

#3

Attorney Docket No. 826.1795

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:

Fumirou ABE, et al.

Application No.:

Group Art Unit: 2175

Filed: February 25, 2002

Examiner: *yaner*

J1050 U.S. PTO  
10/083469  
02/27/02

For: FULL TEXT SEARCH SYSTEM

**SUBMISSION OF CERTIFIED COPY OF PRIOR FOREIGN  
APPLICATION IN ACCORDANCE  
WITH THE REQUIREMENTS OF 37 C.F.R. § 1.55**

Assistant Commissioner for Patents  
Washington, D.C. 20231

Sir:

In accordance with the provisions of 37 C.F.R. § 1.55, the applicant(s) submit(s) herewith a certified copy of the following foreign application:

Japanese Patent Application No. 2001-220256

Filed: July 19, 2001

It is respectfully requested that the applicant(s) be given the benefit of the foreign filing date(s) as evidenced by the certified papers attached hereto, in accordance with the requirements of 35 U.S.C. § 119.

Respectfully submitted,

STAAS & HALSEY LLP

Date: February 25, 2002

By: \_\_\_\_\_

*[Signature]*  
James D. Halsey, Jr.  
Registration No. 22,729

700 11th Street, N.W., Ste. 500  
Washington, D.C. 20001  
(202) 434-1500

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2001年 7月19日

出 願 番 号

Application Number:

特願2001-220256

[ST.10/C]:

[JP2001-220256]

出 願 人

Applicant(s):

富士通株式会社

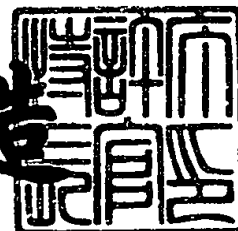


CERTIFIED COPY OF  
PRIORITY DOCUMENT

2002年 1月11日

特許庁長官  
Commissioner,  
Japan Patent Office

及川耕造



出証番号 出証特2001-3113888

【書類名】 特許願

【整理番号】 0151399

【提出日】 平成13年 7月19日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 17/30

【発明の名称】 全文検索システム及び全文検索プログラム

【請求項の数】 5

【発明者】

    【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

    【氏名】 安倍 史郎

【発明者】

    【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

    【氏名】 松浦 正卓

【発明者】

    【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

    【氏名】 田端 洋子

【発明者】

    【住所又は居所】 福岡県福岡市早良区百道浜2丁目2番1号 株式会社富士通九州システムエンジニアリング内

    【氏名】 永田 真彦

【発明者】

    【住所又は居所】 福岡県福岡市早良区百道浜2丁目2番1号 株式会社富士通九州システムエンジニアリング内

    【氏名】 原 泰久

【特許出願人】

    【識別番号】 000005223

【氏名又は名称】 富士通株式会社

【代理人】

【識別番号】 100072590

【弁理士】

【氏名又は名称】 井桁 貞一

【電話番号】 044-754-2462

【手数料の表示】

【予納台帳番号】 011280

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9704486

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 全文検索システム及び全文検索プログラム

【特許請求の範囲】

【請求項 1】 担当する検索担当データの所在と文字列検索条件とを指示されて、検索結果を応答する複数の検索処理装置を有し、検索対象文字列データを文字列レコード群に分割して 1 以上の該検索処理装置に割り当て、与えられた文字列検索条件を検索指示として各該検索処理装置に送信し、各該検索処理装置から検索結果を受信して統合する検索統合手段を設けた全文検索システムであって

上記検索対象文字列データの追加あるいは更新するレコードのみを一時記憶する追加更新一時記憶手段と、

あらかじめ定められた上記いずれか一つの検索処理装置に上記追加更新一時記憶手段に記憶された文字列レコード群を検索担当データとして指示する追加更新レコード検索指示手段と、

を備えることを特徴とする全文検索システム。

【請求項 2】 上記追加更新一時記憶手段に記憶された追加更新レコードに対応する更新前レコードを検索対象文字列データから削除するとともに、上記追加更新レコードを検索対象文字列データに組み入れる更新結果反映手段を備えることを特徴とする請求項 1 記載の全文検索システム。

【請求項 3】 上記検索統合手段が上記複数の検索処理装置に検索指示を送信後、検索処理装置から検索結果を受信した時刻を記憶する検索結果受信時刻記憶手段と、

検索対象文字列データを分割して割り当てられた検索処理装置の内で上記検索結果受信時刻記憶手段により記憶された最初に受信した検索結果受信時刻からあらかじめ設定された時間以内に検索結果を受信できなかった検索処理装置を故障装置として判定する故障検索処理装置判定手段と、

を備えることを特徴とする請求項 1 あるいは請求項 2 のいずれかに記載の全文検索システム。

【請求項 4】 上記故障検索処理装置判定手段が上記検索処理装置の故障を

判定したとき、検索統合手段は上記複数の検索処理装置から送信された検索結果をすべて破棄し、更新結果反映手段を用いて上記追加更新一時記憶手段が記憶した文字列レコード群を検索対象文字列データに組み入れた後、上記故障と判定された検索処理装置および上記追加更新レコード検索指示手段によって検索を指示された検索装置を除いて利用可能な検索処理装置に検索対象文字列データを分割して割り当て、検索指示を行うようにしたことを特徴とする請求項2あるいは請求項3のいずれかに記載の全文検索システム。

【請求項5】 担当する検索担当データの所在と文字列検索条件とを指示されて、検索結果を応答する複数の検索処理装置を有し、検索対象文字列データを文字列レコード群に分割して1以上の該検索処理装置に割り当て、与えられた文字列検索条件を検索指示として各該検索処理装置に送信し、各該検索処理装置から検索結果を受信して統合する検索統合手段を設けた全文検索システムにおいて、コンピュータを、

上記検索対象文字列データの追加あるいは更新するレコードのみを一時記憶する追加更新一時記憶手段、

あらかじめ定められた上記いずれか一つの検索処理装置に上記追加更新一時記憶手段に記憶された文字列レコード群を検索担当データとして指示する追加更新レコード検索指示手段、

として機能させることを特徴とする全文検索プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、複数処理装置を用いて大量データを短時間で検索する文字列照合方式の全文検索システムに関する。ここで、文字列照合方式とは、指定した文字列と検索対象のテキストデータを、検索対象のテキストデータの先頭より後方に向かって逐次照合しながら、指定した文字列が検索対象のテキストデータ中に存在するかを調べる方式である。

【0002】

【従来の技術】

従来にも、特開平8-137734号公報「情報処理システム及びデータベースの分散方法」に、複数の検索処理装置を用いて検索時間の短縮を行う方法として、パソコン等の小規模の情報処理装置をパソコンLANで接続し、データベースを各情報処理装置でもほぼ均一になるようにデータを分散して配置し、データ処理時には各情報処理装置で並列して処理を行ない、処理結果を一つの情報処理装置に統合して表示するようにした方法が提案されている。

## 【 0 0 0 3 】

また、検索対象データを複数の領域に分割して、その分割された領域単位に別の検索処理装置に割当て、それぞれの検索処理装置が同時並行的に検索処理を行う方式が考えられていた。例えば、特開平9-134364号公報「情報検索システム」には、検索処理を行なう複数の検索サーバと、これら複数の検索サーバの動作を管理するための検索管理サーバとを備え、検索管理サーバを、検索対象テキストベースとこのテキストベースについての関連情報とをそれぞれ分割し、このテキストベースの分割部分とその分割部分に対応する関連情報とを組にして複数の検索サーバのうちの一部または全部に割り当てるように構成するとともに、複数の検索サーバを、それぞれ、検索管理サーバによって割り当てられたテキストベースの分割部分に対する情報検索を並列かつ独立に行なうように構成して、情報検索システムに関し、複数のテキストベースを同時に検索する場合でも、特別なハードウェアを必要とせずに、且つ、検索対象テキストベースの大きさにも影響されずに高速に情報検索を行なうようにして同時並行的に検索処理を行う方式が提案されている。

## 【 0 0 0 4 】

## 【発明が解決しようとする課題】

しかし、この上記第1の従来の技術では、データベースの分割において、それぞれを物理的に独立して各装置のファイルとした場合、検索処理装置の増設や検索処理装置の故障による検索処理装置の増減が発生した場合、検索対象データの物理的な再分割が必要となり、その処理のためのシステム停止の時間が長くなることになる。

## 【 0 0 0 5 】

また、上記第2の従来技術の場合、上記のテキスト形式ファイルは、データを追加したり一部のデータを変更すると、ファイル全体の書き換えが行われ、ファイルが巨大な場合、その書き換え処理には大きな時間が必要となる。従って、検索処理とデータの追加更新処理が並行して行われる文字列照合方式の全文検索システムにおいては、この書き換え処理時間の問題を解決しないと実用的なものにならない。また、多くの検索処理装置を用いることにより、検索処理装置のいずれかに障害が発生する可能性はそれだけ高くなることが考えられる。

## 【0006】

したがって、本発明は、検索対象データの更新、追加を検索と並行して行うことが可能で、かつ、いずれかの検索処理装置に障害が発生した場合にも、その障害が発生した検索処理装置を取り除いた縮退運転を迅速に行うことを可能とする手段を提供することを課題とする。

## 【0007】

## 【課題を解決するための手段】

図1には本発明の構成図を示す。図において1は検索対象文字列データである。すなわち、本来の検索を行う検索対象である文字列データを示し、磁気ディスク装置や半導体記憶装置などに格納された論理的にひとつながりの文字列データである。これらは文字列を走査して読み取ることにより識別可能なレコードに区切られており、この文字列レコードがそれぞれまとまった情報として集められ格納されている。検索はこのレコードを構成する文字列情報を検索条件に基づいて照合し満足する文字列レコードを探し出して出力するように行われる。

## 【0008】

4は検索処理装置であり複数台数で構成される。本発明の全文検索システムは上記検索対象文字列データ1をこれら複数の検索処理装置4に分配して検索処理を行うようにするものであり、2は検索対象文字列データ1を文字列レコード群に分割して複数の検索処理装置4に割り当て、与えられた文字列検索条件を検索指示として各検索処理装置4に送信し、各検索処理装置4から検索結果を受信して統合する検索統合手段である。

## 【0009】



通常、検索対象とする文字列データはその内容は固定されたものではなく、新しいデータの追加や更新が必要となる。本発明においてはこのようなデータの追加、更新を検索処理の実行中に並行して実行可能となるなように、上記検索対象文字列データ 1 のレコードを追加更新するとき、不図示の追加更新手段により追加更新したレコードを記憶する追加更新一時記憶手段 3 を備える。これにより、追加、更新時には検索対象文字列データ 1 の内容は変更せず追加更新一時記憶手段 3 に追加、更新データを記憶するようにしている。

## 【 0 0 1 0 】

追加更新されたレコードも検索処理の対象とするために、あらかじめ定められたいずれか一つの検索処理装置 4 に追加更新一時記憶手段 3 に記憶された文字列レコード群を検索担当データとして指示する追加更新レコード検索指示手段 8 を備えるようにしている。

追加更新されたレコードの反映は適時に実施可能とするように、この追加更新一時記憶手段 3 に記憶された更新済みレコードに対応する更新前レコードを検索対象文字列データ 1 から削除するとともに、上記追加更新済みレコードを検索対象文字列データ 1 に組み入れる更新結果反映手段 5 を備えるようにしている。

## 【 0 0 1 1 】

多数の検索処理装置 4 に処理を分担させるとき、その中に故障装置があるときには、その検索処理装置を除外して検索処理を分担して実行する。すなわち、上記検索統合手段 2 が上記複数の検索処理装置 4 に検索指示を送信後、検索対象文字列データ 1 を分割して割り当てられた検索処理装置 4 から検索結果を受信した時刻を記憶する検索結果受信時刻記憶手段 6 と、最初に受信した検索処理装置 4 の検索結果受信時刻から、あらかじめ設定された時間以内に検索結果を受信できなかった検索処理装置 4 を故障装置として判定する故障検索処理装置判定手段 7 とを備える。

## 【 0 0 1 2 】

また、上記追加更新一時記憶手段 3 に記憶されたレコード群を検索担当データとして指示された検索処理装置 4 も記憶されたレコード群が多くなり検索時間が長くなったときに一旦故障装置であると判定されるが、この場合は更新結果反映

手段 5 による追加更新済みレコードの検索対象文字列データ 1 への組み入れを行った後であらためて検索処理を実行する。

## 【 0 0 1 3 】

## 【発明の実施の形態】

本実施の形態においては、パーソナルコンピューター、ワークステーション等の汎用的な目的で使用される計算機上で実行するコンピュータプログラムにより実現する形態を示す。

本発明の全文検索システムは、処理装置、主記憶装置、入出力装置などから構成される計算機上で、コンピュータプログラムを実行して実現される。また、コンピュータプログラムは、フロッピーディスクや CD-ROM 等の可搬型媒体やネットワーク接続された他の計算機の主記憶装置や補助記憶装置等に格納されて提供される。本発明の記録媒体は、上記可搬型媒体、主記憶装置、補助記憶装置に該当する。

## 【 0 0 1 4 】

提供されたコンピュータプログラムは、可搬型媒体から直接計算機の主記憶装置にロードされ、または、補助記憶装置を備えた計算機においては可搬型媒体から一旦補助記憶装置にコピーまたはインストール後に、主記憶装置にロードされて実行する。また、ネットワーク接続された他の装置に格納されて提供された場合も、他の装置からネットワークを経由して受信後に、主記憶装置にロードされ、あるいは補助記憶装置を備える計算機においては補助記憶装置にコピー後に、主記憶装置にロードされて実行するものである。

## 【 0 0 1 5 】

図 2 に本発明の全文検索システム（以後は本システムという）の全体概念を示す。本システムは、複数の端末（端末 A、端末 B、端末 C）から非定期的に発生する検索依頼に対して、検索対象データ（図 1 の検索対象文字列データ 1 に対応する）の中からそれぞれの検索依頼を満足する情報を抽出して、その情報を検索結果としてそれぞれの端末に返す装置である。また、本システムは、検索処理と並行して、端末（端末 D）からのデータの追加更新依頼により、検索対象データへの新しいデータの追加、および検索対象データの変更が行うことができる。

## 【0016】

最初に、検索対象データの内容の例を図3に示す。検索対象データは複数のレコードから構成される。レコードは複数の項目から構成される。項目の中身は文字列である。なお、数値や日付も文字列で表される。レコードはレコード区切符号等の手段により識別できる。レコード内の項目は項目区切符号により区切られており、項目識別符号（以下項目タグと呼ぶ）により一意に特定できる。また、レコードにはレコードを一意に識別できる符号であるレコードNOが付与されている。また、本発明の実施の形態の検索対象データは十分に大量であるとして説明する。

## 【0017】

検索依頼の例を図4に示す。また、図5および図6に端末から検索を行ったときの検索依頼画面の例、検索結果画面の例を示す。この例では、検索依頼は項目タグと検索語を対とした情報の羅列と、検索条件式を持つ。これらの情報は、検索対象レコードの中の項目タグで指定する項目の中に指定検索語が存在し、かつ、検索条件式を満足するレコードを捜して欲しい、と言うことを意味する。

## 【0018】

この例は具体的には、人事情報の検索において、検索条件として、名前と出身校を入力し、関連する情報としてその両方を満足する人事情報を検索するという簡単な例を示している。また、検索依頼画面で名前だけとか名前と出身地とか、入力画面に分かっている情報を入力すると満足する情報が見つかったときにはその関連情報を含めて返すようになっている。このように、入力された情報の組み合わせにより図4に示す検索依頼の例に示す条件式は自動的に作りだされるようになっている。

## 【0019】

このように、本システムは検索依頼を満足するレコードをファイルの中から見つけ出し、そのレコードの一部または全体を端末の画面など検索依頼元に返す。本方式の場合、複数の検索処理装置（図1における検索処理装置4に対応する）を用いて同時並行的に検索処理を行うので、検索処理装置の数に比例して検索時間が短縮される。

## 【 0 0 2 0 】

また、本方式の場合、分割された領域を物理的に独立したファイルとするのではなく、検索対象データを物理的に一つのファイル、あるいは、論理的にひとつながりのファイルに構成し、レコード単位に処理装置の数だけの領域に論理的に分割して、各々の領域を別の検索処理装置に割り当てて検索を行う。これにより、検索処理装置の増減が発生した場合、新たにファイルを検索処理装置の数だけの領域に論理的に再分割を行って、各領域を検索処理装置へ再割当てを行うだけで、速やかに処理を続行することができる。

## 【 0 0 2 1 】

本方式でのファイル形式は、文字情報を単純に並べただけのテキスト形式のファイルとする。しかし、本発明の対象とする全文検索システムの扱う検索対象データは十分に大量であることを前提としているため、検索対象データを一つのファイルにまとめた場合、そのファイルは巨大なものとなる。

上記のテキスト形式ファイルは、データを追加したり一部のデータを変更するとファイル全体の書き換えが行われ、ファイルが巨大な場合、その書き換え処理には大きな時間が必要となる。従って、本発明の対象とする検索処理とデータの追加更新処理が並行して行われる文字列照合方式の全文検索システムにおいては、この書き換え処理時間の問題を解決しないと実用的なものとならない。また、本システムは多くの検索処理装置を用いるため、検索処理装置のいずれかが故障する確率が高くなる。従って、いずれかの検索処理装置に障害が発生した場合に、その障害が発生した検索処理装置を取り除いた縮退運転を迅速に行う必要がある。

## 【 0 0 2 2 】

本発明の実施の形態の全体構成図を図 7 に示す。本システムは、1 つの受付サーバ、複数の検索サーバおよび 1 つのファイルサーバで構成される。受付サーバは端末からの検索依頼を受け付け、それを検索サーバ a ～ e に渡す。検索サーバ a ～ e は、検索依頼を満足するレコードを検索対象ファイルの自分の受け持つ領域の中から見つけ出し、その結果を受付サーバに返す。

## 【 0 0 2 3 】

検索サーバは、検索依頼を満足するレコードを更新ファイルの中から見つけ出し、その結果を受付サーバに返す。受付サーバは、全検索サーバからの回答をとりまとめて端末に返信する。

また、受付サーバは、端末からの検索対象ファイルへのデータの追加更新依頼により、更新ファイルに、新規データの追加登録または変更されたデータの登録を行う。ファイルサーバは、受付サーバまたは検索サーバからの要求により検索対象ファイルおよび更新ファイルのデータの読み出し、および書き込みを行う装置であり、一般的な装置であるので本実施例では説明を割愛する。

#### 【 0 0 2 4 】

次に、受付サーバの詳細について説明する。受付サーバの構成は図 8 の通りである。受付サーバは、情報処理装置、磁気ファイル装置およびプログラム記憶装置で構成される。プログラム記憶装置の構成要素は、制御プログラム、ディレクタプログラム（図 1 における検索統合手段 2 に対応する）、データ更新プログラム、時間監視プログラム（図 1 における検索結果受信時刻記憶手段 6 および故障検索処理装置判定手段 7 に対応する）、検索スレッド、環境テーブル、応答用テーブル、更新ファイル応答用テーブル、検索依頼バッファおよび検索結果テーブルである。

#### 【 0 0 2 5 】

ここで、データ更新プログラムは検索対象の文字列データが格納されているファイルのレコードの更新を行うプログラムである。レコードの更新は検索対象の維持管理を行う受付サーバが行う。

ディレクタプログラムは複数の検索サーバへの検索対象の文字列データを分配したり、端末から受け付けた検索依頼を各検索サーバに受け渡したり、各検索サーバから返される検索結果を統合して依頼元に返すという分割検索を統合する役割を果たすプログラムである。

#### 【 0 0 2 6 】

時間監視プログラムは検索依頼を検索サーバに依頼した後、検索サーバからの応答を監視して、他の検索サーバの応答に比べて極端に時間がかかる検索サーバの異常を検出するものである。

検索スレッドは到着した検索依頼を検索サーバに連続して並行的に送り出すものである。

## 【 0 0 2 7 】

環境テーブル、応答用テーブル、更新ファイル応答用テーブルは後に詳述するが、検索サーバへの検索処理の配分を管理するテーブルである。

検索依頼バッファは受付サーバに端末からの検索依頼を処理待ちの状態として蓄積するバッファであり、各検索サーバに順次検索指示として送信するものである。検索結果テーブルは、各検索サーバが分割された領域の検索対象データを検索した結果を蓄積するもので、全検索サーバの結果が揃ったときこれを統合して依頼元の端末に検索結果を送信することになる。

## 【 0 0 2 8 】

次いで、検索サーバの詳細について説明する。検索サーバの構成を図 9 に示す。検索サーバは、情報処理装置、磁気ファイル装置およびプログラム記憶装置で構成される。プログラム記憶装置の構成要素は、制御プログラム、初期設定プログラム、全文検索プログラム、割当テーブルおよび検索データである。

処理の流れを、事前処理、更新処理、検索処理および領域再割当処理の 4 つに大きく分けて説明する。その 4 つの処理の関係を図 1 0 に示す。事前処理は、システムの起動を行った後、検索依頼またはデータ更新依頼を端末から受け取る前までの処理である。更新処理は、端末からの依頼により新規データの追加登録および変更されたデータの登録を行う処理である。検索処理は、端末から検索依頼を受け取ってから、検索結果を端末に返すまでの処理である。領域再割当処理は、検索処理の過程において検索処理が異常終了した場合に、データ領域の検索サーバへの再割当を行う処理である。

## 【 0 0 2 9 】

まず、事前処理について説明する。事前処理の基本的な機能とテーブルの関連を図 1 1 に示す。図 1 1 には、図 8 および図 9 に示した受付サーバ、検索サーバの構成図のうち事前処理の動作に関わるテーブル類と処理を示し、さらに、検索対象ファイルの各検索サーバでの検索の領域分担の例示、更新ファイルとの関連を示している。

## 【 0 0 3 0 】

図 1 1 に示す検索対象ファイルは検索サーバ a、検索サーバ b、検索サーバ d の 3 台で領域を分担している様子を例示している。各領域はレコードを単位として割り当てられ、例えば、検索サーバ a がレコード NO の 1 から 7 8 4、検索サーバ b が 7 8 5 から 1 5 6 8 のように分担するようにしていることを示している。通常、負荷の均等化を計るためレコード数あるいは文字列の長さ、もしくはサイズの総和がほぼ等しくなるように分割している。また、更新ファイル（図 1 における追加更新一時記憶手段 3 に対応する）は追加された、あるいは更新済のレコードを検索対象ファイルとは別のファイルとして格納するものである。

## 【 0 0 3 1 】

このような内容の図 1 1 を参考にしながら、受付サーバおよび検索サーバの処理の流れを詳細に説明していく。受付サーバにおける動作は、最初に、受付サーバの制御プログラムが、システムの初期化等を行った後、ディレクタプログラムを起動する。ディレクタプログラムの、起動された後の処理の流れを図 1 2 に示す。

## 【 0 0 3 2 】

ディレクタプログラムは、最初にステップ S 1 2 0 1 で更新ファイルのデータを検索対象ファイルに併合（図 1 における更新結果反映手段 5 に対応する）する。更新ファイルは検索対象ファイルの更新されたり追加されたレコードが一時的に格納されるファイルであるので中身がない場合もあるが、通常は前回運用時に蓄積されたレコードが格納されているので、検索を開始あるいは検索の中断等何らかの原因で再開するときにはまず検索対象ファイルに併合するようになっている。

## 【 0 0 3 3 】

このとき、更新ファイルと検索対象ファイルに同一レコード NO のデータが存在した場合、検索対象ファイルのデータを更新ファイルのデータで置き換える。本併合処理の後、更新ファイルのレコードはすべて検索対象ファイルに一本化されるので、更新ファイルのデータは全て消去する。

次に、ステップ S 1 2 0 2 においてディレクタプログラムは環境テーブルを作

成する。図 1 3 に示すように、環境テーブルは検索サーバの一覧であり、検索サーバ名、アドレス（ネットワークアドレス）、状態フラグ、更新ファイル担当フラグから成る。状態フラグとは、その検索サーバの状態を示す。例えば、その検索サーバが使えない状態にあれば「1」とし、使える状態であれば「0」とする。更新ファイル担当フラグは更新ファイルの検索を担当する検索サーバであることを示す。更新ファイル担当フラグ以外の情報は、予め磁気ファイル装置に記録しておき、ディレクタプログラムはそこから情報を得て環境テーブルを設定する。

#### 【 0 0 3 4 】

次に、ステップ S 1 2 0 3 でディレクタプログラムは検索対象ファイルを読んで応答用テーブルを作成する。応答用テーブルは図 1 4 に示すように検索対象ファイル中の全レコードの一覧であり、レコード NO、そのレコードのファイル上の先頭アドレス、およびそのレコードの長さから成る。

次に、ステップ S 1 2 0 4 においてディレクタプログラムは、環境テーブルおよび応答用テーブルを基にして検索サーバ別に、割当テーブルを作成する。割当テーブルは、図 1 5 に示すように各検索サーバの検索の受持ち範囲を割り当てたテーブルであり、環境テーブル（図 1 3 に例示）の状態フラグが使用可能状態「0」である検索サーバの内、一つを除いた全てについてそれぞれ作成する。また、環境テーブル（図 1 3）中の、その残された検索サーバの更新ファイル担当フラグ（図 1 における追加更新レコード検索指示手段 8 に対応する）を ON「1」にする。割当テーブルは個々の検索サーバが検索を受け持つレコードの一覧であり、レコード NO、そのレコードの先頭アドレス、およびそのレコードの長さから成る。個々の検索サーバが受け持つレコードは、応答用テーブル（図 1 4）を見て、個々の割当テーブルのレコードの数またはレコードの長さの総和、が均等になるように、応答用テーブルの先頭から順番に割り当てる。

#### 【 0 0 3 5 】

ステップ S 1 2 0 5 では、ディレクタプログラムは、割当テーブル（図 1 5）を、検索対象ファイル名と共に、それぞれが対応する検索サーバへ通知する。また、環境テーブルの更新ファイル担当フラグが ON「1」の状態にある検索サー



バに対して、中身が空の割当テーブルを更新ファイル名と共に通知し、その後、検索依頼待ち状態となる。以上が図 1 2 の受付サーバのディレクタプログラムの流れに沿った受付サーバの事前処理の動作の説明である。

#### 【 0 0 3 6 】

次に検索サーバにおける事前処理の動作について説明する。最初に、検索サーバの制御プログラムが、システムの初期化を行った後に、初期設定プログラムおよび全文検索プログラムを起動する。

起動された後の初期設定プログラムの処理の流れを図 1 6 に示す。ステップ S 1 6 0 1 で、初期設定プログラムは割当テーブルを受付サーバから受け取り、ステップ S 1 6 0 2 で、割当テーブルに割当レコードがある場合、ステップ S 1 6 0 3 で、その全てのレコード N O のレコードを検索対象ファイルよりプログラム記憶装置に読み込み検索データ（図 1 7）を作成する。

#### 【 0 0 3 7 】

割当テーブルに割当レコードがない場合は、更新ファイル担当の検索サーバであるので何もしないでそのまま終わる。以後、全文検索プログラムは起動された後、検索依頼待ち状態となる。

次に、上記事前処理の説明に続いて、図 1 0 の中に示した更新処理について説明する。更新処理の基本的な機能とテーブルの関連を図 1 8 に示す。図 1 8 を参考にしながら、更新処理における受付サーバの処理の流れを詳細に説明する。

#### 【 0 0 3 8 】

受付サーバの制御プログラムは、システムの初期化等を行った後、ディレクタプログラムと共にデータ更新プログラムを起動する。起動された後のデータ更新プログラムの処理の流れを図 1 9 に示す。

データ更新プログラムは、起動された後、プログラムの初期化等を行い、データ更新依頼待ちとなる。データ更新依頼を受け付けた後、ステップ S 1 9 0 1 において、送られてきたデータが新規の追加登録データであれば、従来検索対象ファイルおよび更新ファイルに存在した最大レコード N O + 1 のレコード N O を付与して、更新ファイルの最後に記録する。送られてきたデータが変更されたデータであれば、そのまま更新ファイルの最後に記録する。もし、送られてきた変更

データと同じレコードNOのデータが更新ファイルに存在した場合は、存在していた更新ファイルのデータを削除する。更新ファイルへのデータ登録後、ステップS1902で、更新ファイル応答用テーブル(図20)を作成して、再びデータ更新依頼待ちとなる。

#### 【0039】

次に、図10に示した処理の全体の流れの中の検索処理について説明する。検索処理の基本的な機能とテーブルの関連を図21に示す。図21を参考にしながら、検索処理における受付サーバおよび検索サーバの処理の流れを詳細に説明する。ここに示すように、端末から検索依頼を受け取ると、事前処理でディレクタプログラムが設定したように、検索対象ファイルの各割り当てられた領域のレコードが検索サーバa、b、dに、更新ファイルのレコードが検索サーバeに検索を担当させるようにして、検索依頼が各検索サーバに送信される。

#### 【0040】

図12の処理の流れの最後で、ディレクタプログラムは事前処理を終了すると検索依頼待ちの状態にある。その後の処理の流れを図22に示す。

端末からの検索依頼は検索依頼バッファ(図23)に格納される。ステップS2201では、端末から検索依頼があると、ディレクタプログラムは検索依頼バッファから検索依頼を取り出し、まず、検索結果テーブル(図24)を検索依頼単位に作成する。検索結果テーブルには環境テーブル(図13)の使える状態にある検索サーバ名の一覧を作成し、環境テーブルの更新ファイル担当フラグを転記し、更に、処理結果フラグ、処理完了時刻およびヒットレコードNOの欄を設ける。

#### 【0041】

次に、ステップS2202で、ディレクタプログラムは、使える状態にある全検索サーバに対応した検索スレッドを起動する。検索スレッドには、起動と同時に、検索依頼(図4)、環境テーブル(図13)中の担当する検索サーバのアドレスおよび検索結果テーブル(図24)中の自分が回答を書き込む欄を知らせる。検索スレッドは検索依頼別に、使える状態にある検索サーバの数だけ作成されることになる。検索スレッドを起動した後、ディレクタプログラムはステップS

2 2 0 3 で検索結果通知待ちの状態になるが、以降の処理は後述する。

【 0 0 4 2 】

検索スレッドの処理の流れを図 2 5 に示す。検索スレッドは起動された後、検索依頼を受け取り、ステップ S 2 5 0 1 で、それを自分の担当する検索サーバの全文検索プログラムに通知する。ステップ S 2 5 0 2 で検索結果を待ち、ステップ S 2 5 0 3 で検索結果の確認をして、検索サーバから検索結果の回答としてヒットレコード NO ( 図 2 6 ) を受け取ったら、ステップ S 2 5 0 4 で、その回答を検索結果テーブル ( 図 2 7 ) のヒットレコード NO 欄に書き込み、更に、ステップ S 2 5 0 5 で、その時の時刻を処理完了時刻欄に書き込み、ステップ S 2 5 0 6 で処理結果フラグを正常終了状態「1」にする。もし検索サーバから異常終了の回答があったら、ステップ S 2 5 0 6 で処理結果フラグ欄を異常状態「2」にする。

【 0 0 4 3 】

次に、検索スレッドから検索依頼を受け取った後の全文検索プログラムの処理の流れを図 2 8 に示す。全文検索プログラムは、受け取った検索依頼を満足するレコードを、ステップ S 2 8 0 1 で割当テーブル ( 図 1 5 ) にレコードがある場合はステップ S 2 8 0 2 で、プログラム記憶装置上の検索データ ( 図 1 7 ) の中から探し出し、ステップ S 2 8 0 1 で割当テーブルが空の場合はステップ S 2 8 0 3 で更新ファイルから探し出す。次に、ステップ S 2 8 0 4 で探し出したレコードのレコード NO をヒットレコード NO ( 図 2 6 ) として、依頼のあった検索スレッドに回答する。もし、検索過程で異常を検出した場合は、異常終了のメッセージを依頼のあった検索スレッドに回答する。

【 0 0 4 4 】

時間監視プログラムの処理の流れを図 2 9 に示す。時間監視プログラムは一定時間間隔で処理を開始する。処理開始後、全ての検索結果テーブル ( 図 2 7 ) について以下の処理を行う。

まず、ステップ S 2 9 0 1 で、処理結果フラグが異常終了「2」となっている検索サーバがあれば、ステップ S 2 9 0 6 でディレクタプログラムに処理結果通知として異常終了の通知を行う。ステップ S 2 9 0 2 で全処理結果フラグが正常

終了「1」となっていた場合は、ステップS2904でディレクタプログラムに処理結果通知として正常終了を通知する。処理結果フラグが未入力、すなわち処理未完了の検索サーバがあれば、ステップS2903で更新ファイル担当以外の検索サーバの処理完了時刻の内、最も早い時刻より一定時間を経過していた場合は、ステップS2905で処理結果フラグが未入力で、かつ更新ファイル担当以外の検索サーバの検索結果テーブル(図27)の処理結果フラグを異常終了「2」にした後、ステップS2906でディレクタプログラムに処理結果通知として異常終了の通知を行う。

## 【0045】

図22に戻って説明する。ディレクタプログラムは検索スレッドの起動後、ステップS2204で時間監視プログラムからの処理結果通知を受け取る。

処理結果通知の内容が正常終了であった場合はステップS2205で次の処理を行う。更新ファイル担当フラグがOFFすなわち「0」の検索サーバについては、応答用テーブル(図14)から、ヒットレコードNOに記載されているレコードNOに対応するレコードの先頭アドレスおよびレコードの長さを取り出して、それに該当するレコードの内容を検索対象ファイルより読み込む。更新ファイル担当フラグがONすなわち「1」の検索サーバについては更新ファイル応答用テーブル(図20)より、ヒットレコードNOに記載されているレコードNOに対応するレコードの先頭アドレスおよびレコードの長さを取り出して、それに該当するレコードの内容を更新ファイルより読み込む。但し、応答用テーブル(図14)と更新ファイル応答用テーブル(図20)に同一のレコードNOのデータがある場合は、応答用テーブル(図14)のレコードNOのデータを無視する。最後に、ステップS2206で読み込んだレコードの内容を所定の形式に編集して端末に送信し、検索依頼バッファ中の当該端末の検索依頼データを消去する。

## 【0046】

ステップS2204で処理結果通知の内容が異常終了であった場合は、領域再割当処理を行う。ディレクタプログラムの領域再割当処理の処理の流れを図30に示す。領域再割当処理では、まず、ステップS3001で全検索スレッドの停

止を行う。次にステップ S 3 0 0 2 で検索結果テーブル(図 2 7)の処理結果フラグが異常「2」を示している検索サーバについて、環境テーブル(図 3 1)の状態フラグを使用不可状態「1」にした後、全ての検索結果テーブルを消去する。

#### 【0047】

次に、ステップ S 3 0 0 3 で更新ファイルのデータを検索対象ファイルに併合する。このとき、更新ファイルと検索対象ファイルに同一レコード NO のデータが存在した場合、検索対象ファイルのデータを更新ファイルのデータで置き換える。本併合処理の後、更新ファイルのデータは全て消去する。

その後、ステップ S 3 0 0 4 で応答用テーブル(図 1 4)の再創生を行う。

#### 【0048】

次に、ステップ S 3 0 0 5 で応答用テーブル(図 1 4)と環境テーブル(図 3 1)を用いて、図 3 2 に示すように、各検索サーバの検索受持ち範囲の再割当を行う。例えば、図 3 1 に示すように、検索サーバ d が使用不可であることが判明した場合には、図 3 2 に示すように、更新ファイルの担当である検索サーバ e を除いた検索サーバ a、b、d の 3 つに割り当てていた検索対象ファイルを、検索サーバ a、b の 2 つに割り当て直す。ステップ S 3 0 0 6 でその再割当結果の割当テーブル(図 3 3 に形式を示す)を全検索サーバに通知する。

#### 【0049】

検索サーバの初期化プログラムはその通知を受け取って、事前処理と同じ処理(図 1 6)を行う。

以上の処理の後、ディレクタプログラムは、検索依頼バッファにある検索依頼を取り出して、検索処理(図 2 2)を再開する。

以上で一つ目の実施例の説明を終える。

#### 【0050】

二つ目の実施例について説明する。一つ目の実施例は、受付サーバおよび検索対象ファイルが一つの場合の事例である。二つ目の実施例では、受付サーバと検索対象ファイルが組となり複数存在する場合について説明する。二つ目の実施例は、検索依頼者が、地域的に分散して存在する複数の一つ目の実施例のシステム

に対して、横断的な検索依頼を行いたい場合に有効である。

#### 【 0 0 5 1 】

二つ目の実施例の全体構成図を図 3 4 に示す。検索対象ファイルが複数個ある場合、それぞれの検索対象ファイルに対応して一つ目の実施例と同じシステムを作成する。これにより、受付サーバが複数個作成されることになる。これらの受付サーバと端末の間に一つのジェネラル受付サーバを設ける。各端末は、ジェネラル受付サーバに対して検索依頼を行い、ジェネラル受付サーバから検索結果を受け取る。このようなケースにおいては、検索対象ファイルの更新、維持の処理は各管理単位の受付サーバ単位に実施されるので、ジェネラル受付サーバを経由してアクセスする端末からのデータ更新依頼を許可しないのが一般的であり、ジェネラル受付サーバは、端末からのデータ更新依頼を受付ないものとする。

#### 【 0 0 5 2 】

ジェネラル受付サーバ以外は全て一つ目の実施例と同じであるので、ジェネラル受付サーバについてのみ詳細を説明する。ジェネラル受付サーバの構成は図 3 5 のとおりである。ジェネラル受付サーバは、情報処理装置、磁気ファイル装置およびプログラム記憶装置で構成される。プログラム記憶装置の構成要素は、制御プログラム、ディレクタプログラム、環境テーブル G、検索結果テーブル G および受付スレッドである。

#### 【 0 0 5 3 】

二つ目の実施例の基本的な機能とテーブルの関連を図 3 6 に示す。図 3 6 を参考にしながら、ジェネラル受付サーバの処理の流れを詳細に説明する。最初に、制御プログラムがシステムの初期化を行った後、ディレクタプログラムを起動する。

ディレクタプログラムは、起動された後、環境テーブル G（図 3 7）を作成し、その後は検索依頼待ちとなる。環境テーブル G は受付サーバの一覧であり、受付サーバ名、アドレス、状態フラグから成る。状態フラグとは、その受付サーバの状態を示す。例えば、その受付サーバが使えない状態にあれば「1」とし、使える状態であれば「0」とする。これらの情報は予め磁気ファイル装置に記録しておき、ディレクタプログラムはそこから情報を得る。

## 【 0 0 5 4 】

検索依頼を受け取った後のディレクタプログラムの処理の流れを図 3 8 に示す。ディレクタプログラムは、検索依頼を受け取った後、ステップ S 3 8 0 1 で環境テーブルに基づき、検索結果テーブル G (図 3 9) を作成する。検索結果テーブル G は使用可能状態にある受付サーバの一覧であり、処理結果フラグ欄及び検索結果欄を持つ。

## 【 0 0 5 5 】

次に、ディレクタプログラムは、ステップ S 3 8 0 2 で使える状態にある全受付サーバに対応した受付スレッドを起動する。受付スレッドには、起動と同時に、検索依頼、担当する受付サーバのアドレスおよび検索結果テーブル中の自分が回答を書き込む欄を知らせる。受付スレッドは検索依頼別に、使える状態にある受付サーバの数だけ作成されることになる。受付スレッドを起動した後、ディレクタプログラムはステップ S 3 8 0 3 で検索結果通知待ちの状態になるが、以降の処理は後述する。

## 【 0 0 5 6 】

受付スレッドの処理の流れを図 4 0 に示す。受付スレッドは起動された後、ステップ S 4 0 0 1 で検索依頼を受け取り、それをステップ S 4 0 0 2 で自分の担当する受付サーバに送信する。ステップ S 4 0 0 3 で受付サーバから検索結果 (図 4 1) を受け取ったら、ステップ S 4 0 0 4 で、それを検索結果テーブル G (図 4 2) の検索結果欄に書き込むと同時に処理結果フラグを正常終了状態 (例えば「1」) にする。ステップ S 4 0 0 3 で、もし一定時間内に受付サーバから回答がなかった場合は、ステップ S 4 0 0 5 で受付サーバに異常があったと見做して、処理結果フラグ欄を異常状態 (例えば「2」) にする。

## 【 0 0 5 7 】

図 3 8 に戻って説明する。ディレクタプログラムは受付スレッドの起動後、検索依頼別に全受付スレッドの処理完了待ちを管理しておき、ステップ S 3 8 0 4 で処理完了となった時点で検索結果テーブル G (図 4 2) の処理結果フラグを確認する。その結果、ステップ S 3 8 0 6、ステップ S 3 8 0 7 で全ての検索結果が正常である受付サーバの検索結果を所定の形式に編集して端末に送信する。

## 【 0 0 5 8 】

処理結果フラグの確認の結果、いずれかの処理結果フラグが異常を示していた場合は、ステップ S 3 8 0 5 で環境テーブル G (図 4 3) の該当する受付サーバの状態フラグを使用不可状態にする。以後に端末から受け取った検索依頼については、使用可能な受付サーバに対してのみ検索依頼を行う。以上で二つ目の実施例の説明を終える。

## 【 0 0 5 9 】

ここで、レコード番号の扱いについて補足して説明しておく。上記実施の形態において、検索対象文字列データの中のレコードの構成としてレコード NO を保持するようにしているが、検索条件としては文字列を比較することにより行うので、レコード番号を持つ必要がなくレコードの並びとその分割を説明するために例示したものである。

## 【 0 0 6 0 】

実際、レコード NO は各検索処理装置に検索対象文字列データの検索担当範囲としてレコードを配分するとき受付サーバが応答用テーブルを作成するときに、検索対象文字列データの先頭から順次文字列を読み込むのでレコードの先頭が分かればテーブルに順次レコード番号を採番して記録し、そのファイル (検索対象文字列データ) 上の位置、レコードの長さを記録できる。

## 【 0 0 6 1 】

また、このように、受付サーバは上記応答用テーブルを保持しているので、各検索処理装置への担当領域の配分、再配分を連続したレコード NO を用いて容易に行うことができるようになっている。

また、上記実施の形態の説明においては検索中の検索対象文字列データの保守作業として追加、更新を取り上げたが、更新ファイル応答用テーブル図 2 0 には図示していないが、ここに従来技術により用いられるように削除フラグを設けて、追加更新だけではなく検索処理中に検索対象文字列データの削除も行えるようにすることができる。

## 【 0 0 6 2 】

また、更新ファイルとそれぞれの検索処理装置が分担して検索する検索対象文



文字列データの各領域の両方で見つかった同一レコードNOのレコードの判定により更新されたデータを優先するようにしたが、これも、応答用テーブルなどテーブルへの更新フラグ記録を行って判別してもよい。

これは、更新結果反映手段5による追加、更新、削除などの検索対象文字列データへの反映を行う方法にも適用してもよい。

#### 【 0 0 6 3 】

なお、上記各検索サーバは担当する検索対象のデータの扱いは、従来は、実用的にはファイル上での検索を行うことが多かったが、本発明の実施の形態においては、担当領域のデータをいったん各検索サーバのメモリ上に展開して、文字列の検索を行うようにしている。ただし、更新ファイルについては、更新ファイルに記憶されるレコードの数が少ないときには従来通りファイル上の検索を行ってもよく、上記説明した実施の形態においては、メモリ上への展開は行っていない。更新の頻度が多い場合には上記他の検索サーバと同様に更新ファイル担当の検索サーバに付いてもプログラム記憶装置に検索データを展開して文字列の検索を行ってもよい。

#### 【 0 0 6 4 】

本発明の実施の形態では、各検索処理装置はサーバであり受付サーバとの間はネットワークで接続されているとしたが、検索の受付をする装置と検索をする装置がその他高速リンク等の接続方法を用いても同様の効果が期待できる。

（付記1）担当する検索担当データの所在と文字列検索条件とを指示されて、検索結果を応答する複数の検索処理装置を有し、検索対象文字列データを文字列レコード群に分割して1以上の該検索処理装置に割り当て、与えられた文字列検索条件を検索指示として各該検索処理装置に送信し、各該検索処理装置から検索結果を受信して統合する検索統合手段を設けた全文検索システムであって、

上記検索対象文字列データの追加あるいは更新するレコードのみを一時記憶する追加更新一時記憶手段と、

あらかじめ定められた上記いずれか一つの検索処理装置に上記追加更新一時記憶手段に記憶された文字列レコード群を検索担当データとして指示する追加更新レコード検索指示手段と、

を備えることを特徴とする全文検索システム。

【 0 0 6 5 】

（付記 2）上記追加更新一時記憶手段に記憶された追加更新レコードに対応する更新前レコードを検索対象文字列データから削除するとともに、上記追加更新レコードを検索対象文字列データに組み入れる更新結果反映手段を備えることを特徴とする付記 1 記載の全文検索システム。

（付記 3）上記検索統合手段が上記複数の検索処理装置に検索指示を送信後、検索処理装置から検索結果を受信した時刻を記憶する検索結果受信時刻記憶手段と、

検索対象文字列データを分割して割り当てられた検索処理装置の内で上記検索結果受信時刻記憶手段により記憶された最初に受信した検索結果受信時刻からあらかじめ設定された時間以内に検索結果を受信できなかった検索処理装置を故障装置として判定する故障検索処理装置判定手段と、

を備えることを特徴とする付記 1 あるいは付記 2 のいずれかに記載の全文検索システム。

【 0 0 6 6 】

（付記 4）上記故障検索処理装置判定手段が上記検索処理装置の故障を判定したとき、検索統合手段は上記複数の検索処理装置から送信された検索結果をすべて破棄し、更新結果反映手段を用いて上記追加更新一時記憶手段が記憶した文字列レコード群を検索対象文字列データに組み入れた後、上記故障と判定された検索処理装置および上記追加更新レコード検索指示手段によって検索を指示された検索装置を除いて利用可能な検索処理装置に検索対象文字列データを分割して割り当て、検索指示を行うようにしたことを特徴とする付記 2 あるいは付記 3 のいずれかに記載の全文検索システム。

【 0 0 6 7 】

（付記 5）担当する検索担当データの所在と文字列検索条件とを指示されて、検索結果を応答する複数の検索処理装置を有し、検索対象文字列データを文字列レコード群に分割して 1 以上の該検索処理装置に割り当て、与えられた文字列検索条件を検索指示として各該検索処理装置に送信し、各該検索処理装置から検索

結果を受信して統合する検索統合手段を設けた全文検索システムにおいて、コンピュータを、

上記検索対象文字列データの追加あるいは更新するレコードのみを一時記憶する追加更新一時記憶手段、

あらかじめ定められた上記いずれか一つの検索処理装置に上記追加更新一時記憶手段に記憶された文字列レコード群を検索担当データとして指示する追加更新レコード検索指示手段、

として機能させることを特徴とする全文検索プログラム。

【0068】

(付記6) 担当する検索担当データの所在と文字列検索条件とを指示されて、検索結果を応答する複数の検索処理装置を有し、検索対象文字列データを文字列レコード群に分割して1以上の該検索処理装置に割り当て、与えられた文字列検索条件を検索指示として各該検索処理装置に送信し、各該検索処理装置から検索結果を受信して統合する検索統合手段を設けた全文検索システムにおいて、コンピュータを、

上記検索対象文字列データの追加あるいは更新するレコードのみを一時記憶する追加更新一時記憶手段、

あらかじめ定められた上記いずれか一つの検索処理装置に上記追加更新一時記憶手段に記憶された文字列レコード群を検索担当データとして指示する追加更新レコード検索指示手段、

として機能させるプログラムを記録したコンピュータ読み取り可能な記録媒体。

【0069】

(付記7) 担当する検索担当データの所在と文字列検索条件とを指示されて、検索結果を応答する複数の検索処理サーバとネットワークで接続され、検索対象文字列データを文字列レコード群に分割して1以上の該検索処理サーバに割り当て、与えられた文字列検索条件を検索指示として各該検索処理サーバに送信し、各該検索処理サーバから検索結果を受信して統合する検索統合サーバであって、

上記検索対象文字列データの追加あるいは更新するレコードのみを一時記憶する追加更新一時記憶手段と、

あらかじめ定められた上記いずれか一つの検索処理サーバに上記追加更新一時記憶手段に記憶された文字列レコード群を検索担当データとして指示する追加更新レコード検索指示手段と、

を備えることを特徴とする検索統合サーバ。

#### 【0070】

(付記8) 担当する検索担当データの所在と文字列検索条件とを指示されて、検索結果を応答する複数の検索処理装置を有し、検索対象文字列データを文字列レコード群に分割して1以上の検索処理装置に割り当て、与えられた文字列検索条件を検索指示として各検索処理装置に送信し、各検索処理装置から検索結果を受信して統合する検索統合手段を設けた全文検索システムにおいて、上記検索対象文字列データのレコードを追加あるいは更新するとき、

追加更新するレコードのみを一時記憶するステップ、

あらかじめ定められた上記いずれか一つの検索処理装置に上記記憶された追加更新する文字列レコード群を検索担当データとして指示するステップ、  
とを備えることを特徴とする全文検索中のレコード追加更新方法。

#### 【0071】

##### 【発明の効果】

以上の説明から明らかなように、本発明によれば、レコード区切りのある文字列の並びで構成される大量データから文字列照合方式により与えられた条件の文字列レコードを検索するシステムにおいて、検索処理を行っているときにも並行して、検索対象となる大量データの追加、更新を行うことができるので、大量データの検索システムの運用を容易にするとともに、利用者にとっても検索対象データの時々刻々、追加、更新された最新の情報を検索できるので、検索データの品質向上、信頼性の向上が期待できる。また、本発明の、大量データの検索を多数の検索処理装置を同時に動作させることによって検索速度向上をはかる仕組みにおいても、各検索処理装置が分担する検索対象データの割当も短時間で変更することができるので、もしも一部検索処理装置の不具合が発生しても、検索処理装置の台数を減縮した構成で処理をほとんど中断せずに運転できるので検索システムの全体としての信頼性向上と、稼働率の向上が図れる。

【図面の簡単な説明】

- 【図 1】 本発明の構成図
- 【図 2】 全文検索システムの全体概念
- 【図 3】 検索対象データの例
- 【図 4】 検索依頼の例
- 【図 5】 検索依頼画面の例
- 【図 6】 検索結果画面の例
- 【図 7】 全文検索システムの全体構成図
- 【図 8】 受付サーバの構成図
- 【図 9】 検索サーバの構成図
- 【図 10】 処理の全体の流れ
- 【図 11】 事前処理のブロック図
- 【図 12】 ディレクタプログラムの流れ図 (1)
- 【図 13】 環境テーブルの例 (1)
- 【図 14】 応答用テーブルの例
- 【図 15】 検索サーバ a の割当テーブルの例 (1)
- 【図 16】 検索サーバの初期設定プログラムの流れ図
- 【図 17】 検索データの例
- 【図 18】 更新処理のブロック図
- 【図 19】 データ更新プログラムの流れ図
- 【図 20】 更新ファイル応答用テーブルの例
- 【図 21】 検索処理のブロック図
- 【図 22】 ディレクタプログラムの流れ図 (2)
- 【図 23】 検索依頼バッファ
- 【図 24】 検索結果テーブルの例 (1)
- 【図 25】 検索スレッドの流れ図
- 【図 26】 検索結果の例 (1)
- 【図 27】 検索結果テーブルの例 (2)
- 【図 28】 全文検索プログラムの流れ図

- 【図 2 9】 時間監視プログラムの処理の流れ
- 【図 3 0】 ディレクタプログラムの流れ図(3)
- 【図 3 1】 環境テーブルの例(2)
- 【図 3 2】 検索対象領域の再割当
- 【図 3 3】 検索サーバ a の割当テーブルの例 ( 2 )
- 【図 3 4】 二つ目の実施例の全体構成図
- 【図 3 5】 ジェネラル受付サーバの構成図
- 【図 3 6】 二つ目の実施例の処理ブロック図
- 【図 3 7】 環境テーブル G の例(1)
- 【図 3 8】 ジェネラル受付サーバのディレクタプログラムの流れ図(2)
- 【図 3 9】 検索結果テーブル G の例 ( 1 )
- 【図 4 0】 受付スレッドの流れ図
- 【図 4 1】 検索結果の例(2)
- 【図 4 2】 検索結果テーブル G の例(2)
- 【図 4 3】 環境テーブル G の例(2)
- 【符号の説明】

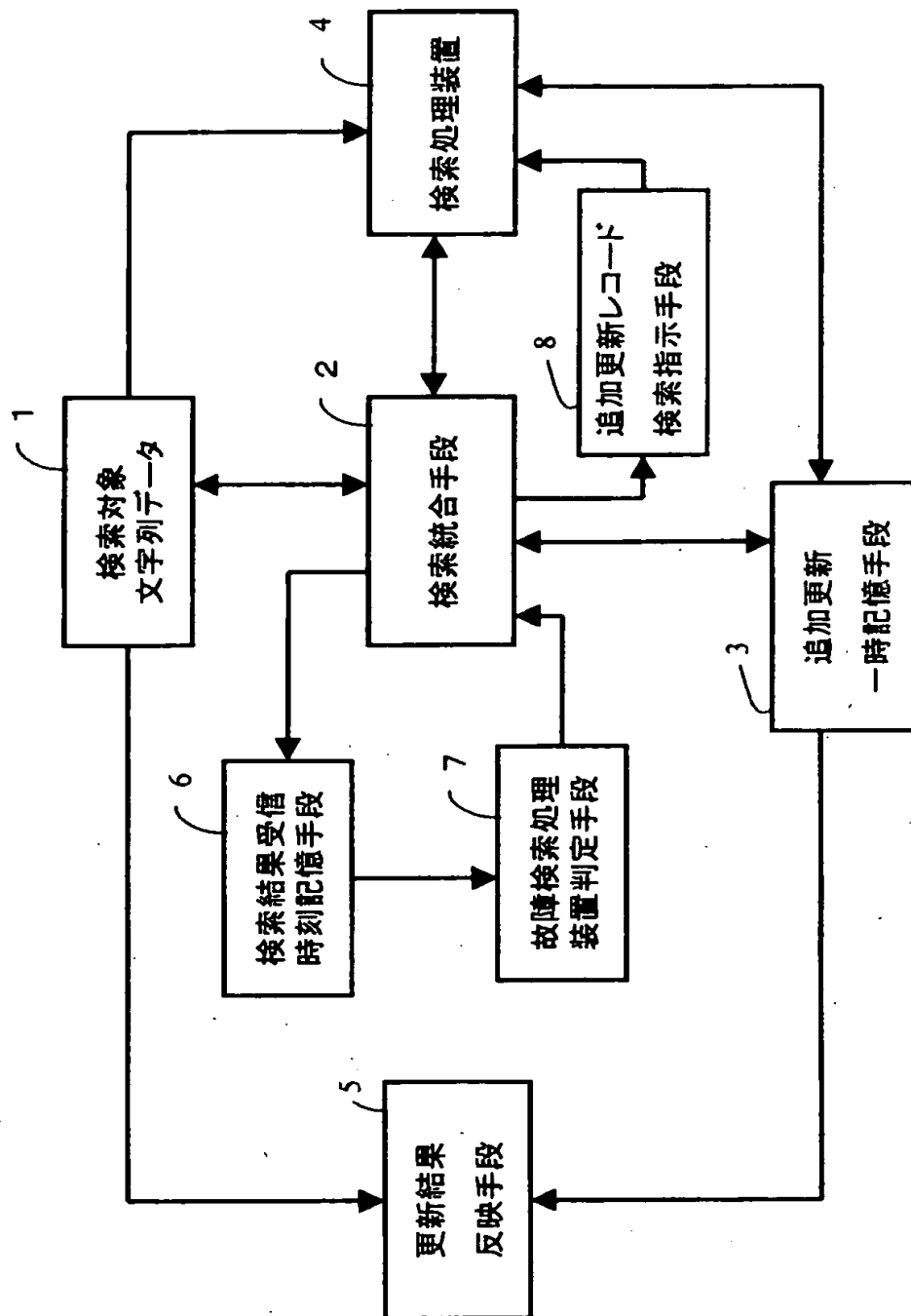
- 1 検索対象文字列データ
- 2 検索統合手段
- 3 追加更新一時記憶手段
- 4 検索処理装置
- 5 更新結果反映手段
- 6 検索結果受信時刻記憶手段
- 7 故障検索処理装置判定手段
- 8 追加更新レコード検索指示手段

【書類名】

図面

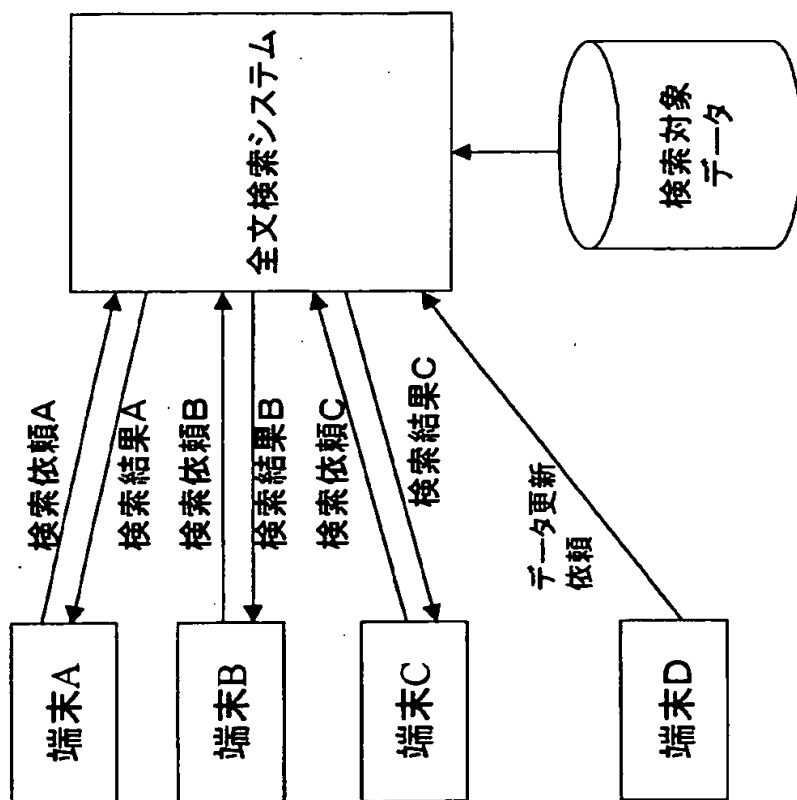
【図 1】

本発明の構成図



【図2】

全文検索システムの全体概念





【図3】

# 検索対象データの例

[K]<NO>1[K]<名前>松浦正卓[K]<出身校>AA大学[K]<出身地>神奈川県[K]<性別>男[K][R]  
 [K]<NO>2[K]<名前>田端洋子[K]<出身校>BB大学[K]<出身地>石川県[K]<性別>女[K][R]  
 [K]<NO>3[K]<名前>安倍史郎[K]<出身校>CC大学[K]<出身地>大分県[K]<性別>男[K][R]

(注) [R]; レコード区切り符号, [K]; 項目区切り符号、<文字列>; 項目識別符号(項目タグ)

【図 4】

検索依頼の例

	項目タグ	検索語
A	<名前>	松浦正卓
B	<出身校>	AA大学
検索条件式	A and B	

【図 5】

検索依頼画面の例

人事情報検索条件

名前

松浦正卓

出身校

AA大学

出身地

▼

性別

▼

【図6】

検索結果画面の例

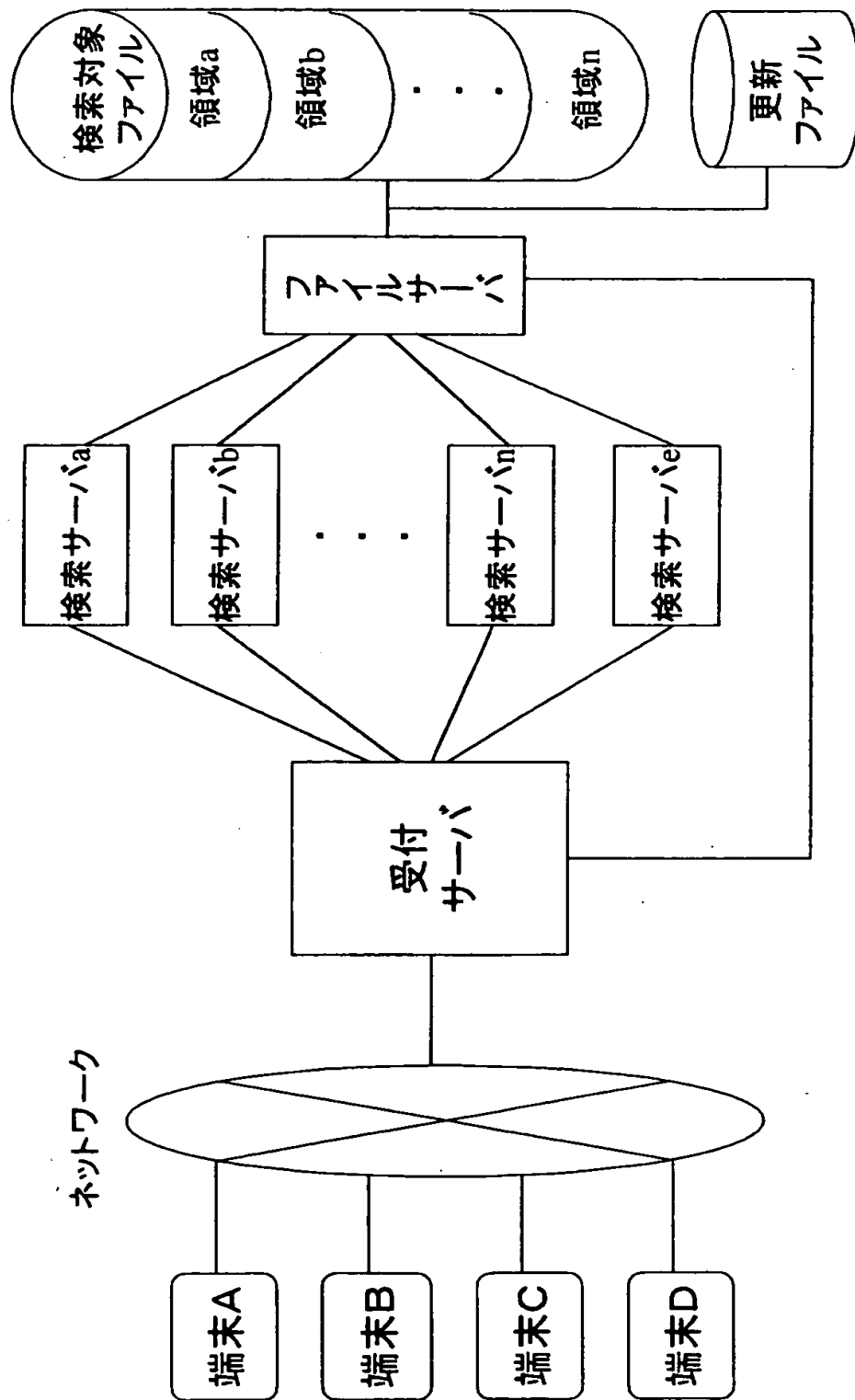
人事情報検索結果

1/1

NO	1
名前	松浦正卓
出身校	AA大学
出身地	神奈川県
性別	男

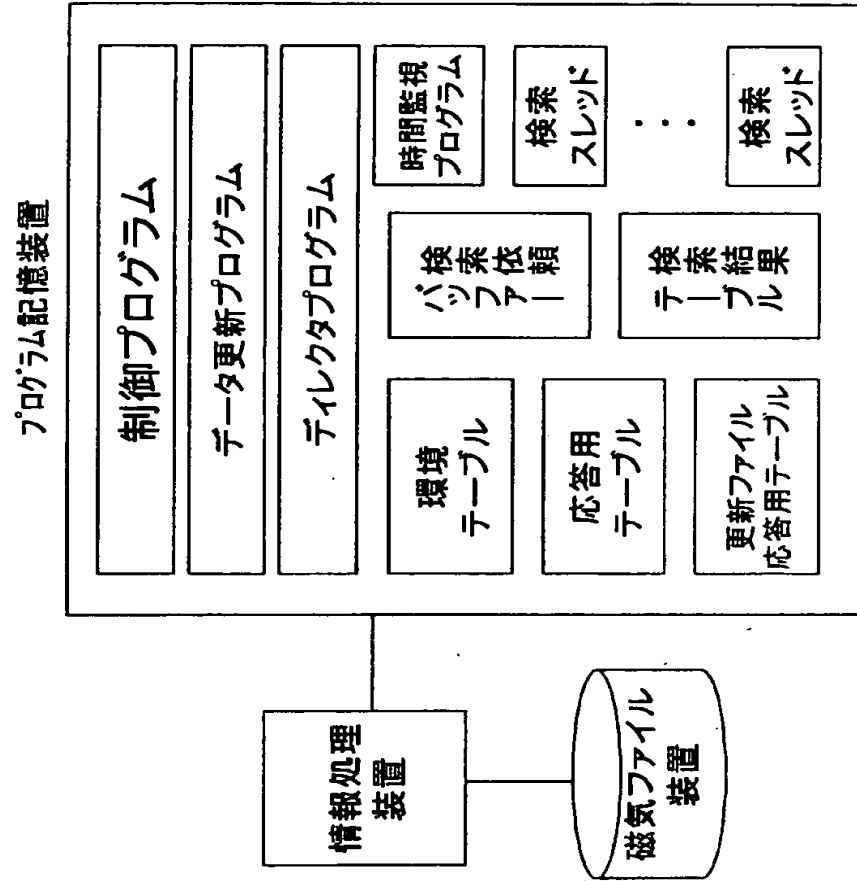
【図7】

全文検索システムの全体構成図



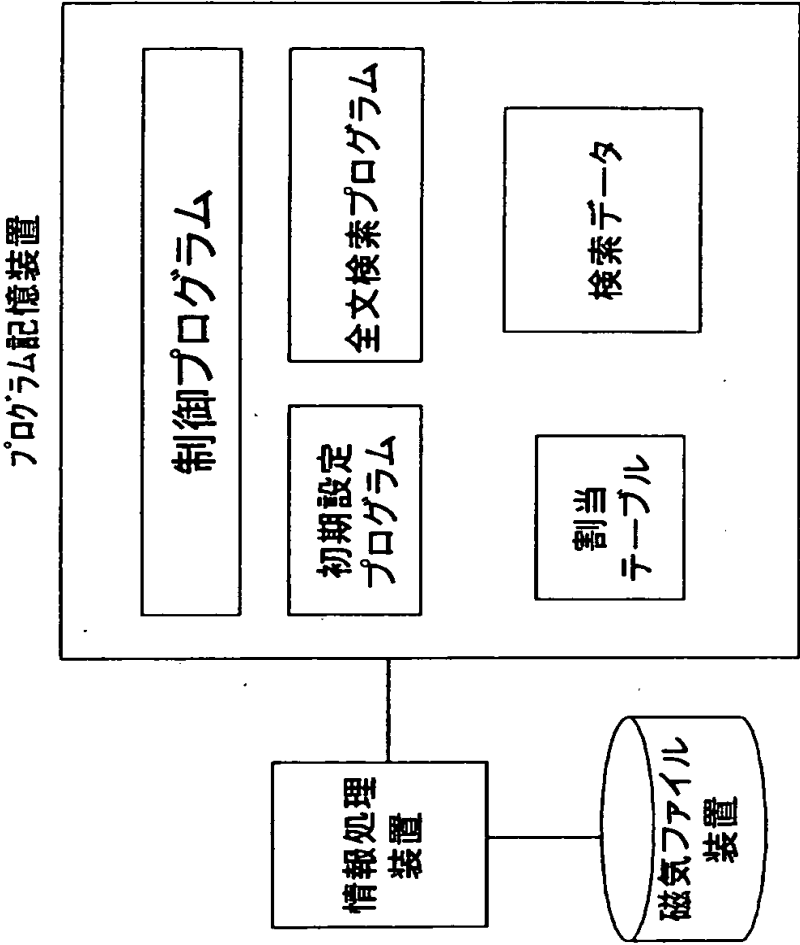
【図 8】

受付サーバの構成図



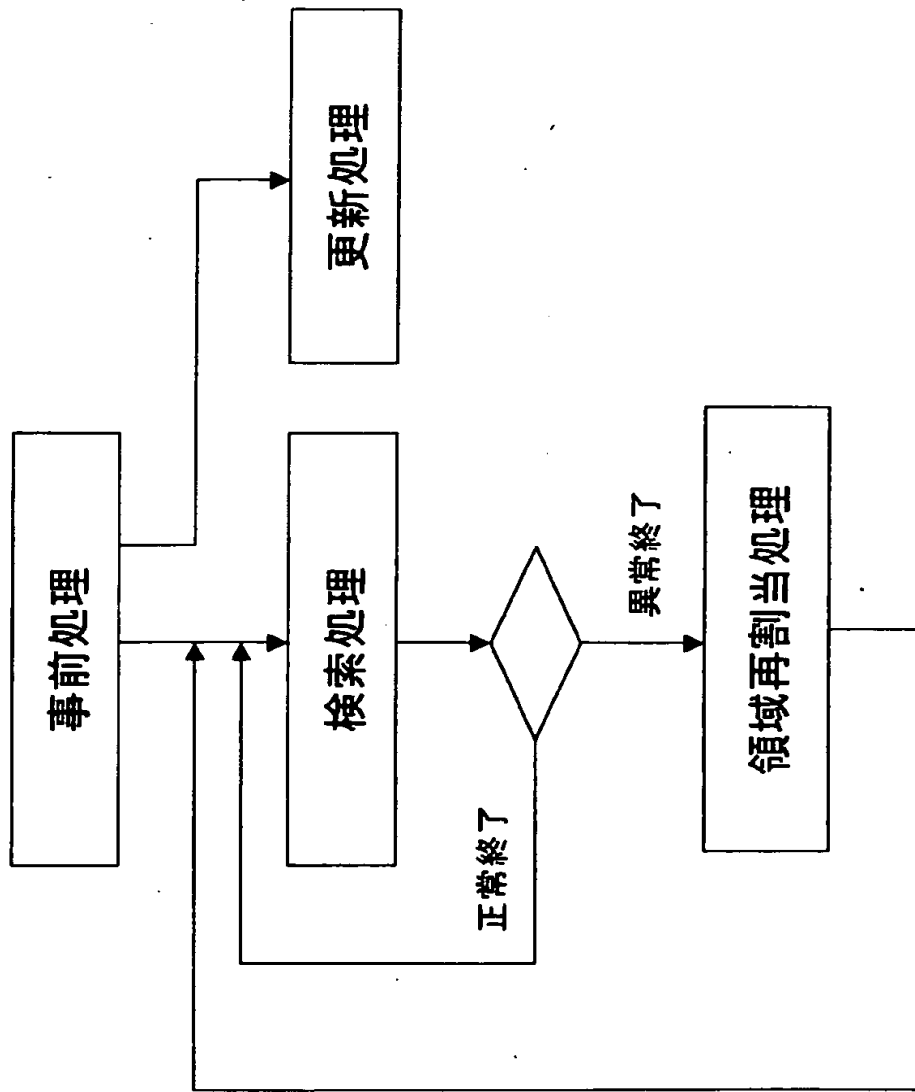
【図9】

検索サーバの構成図



【図 1 0】

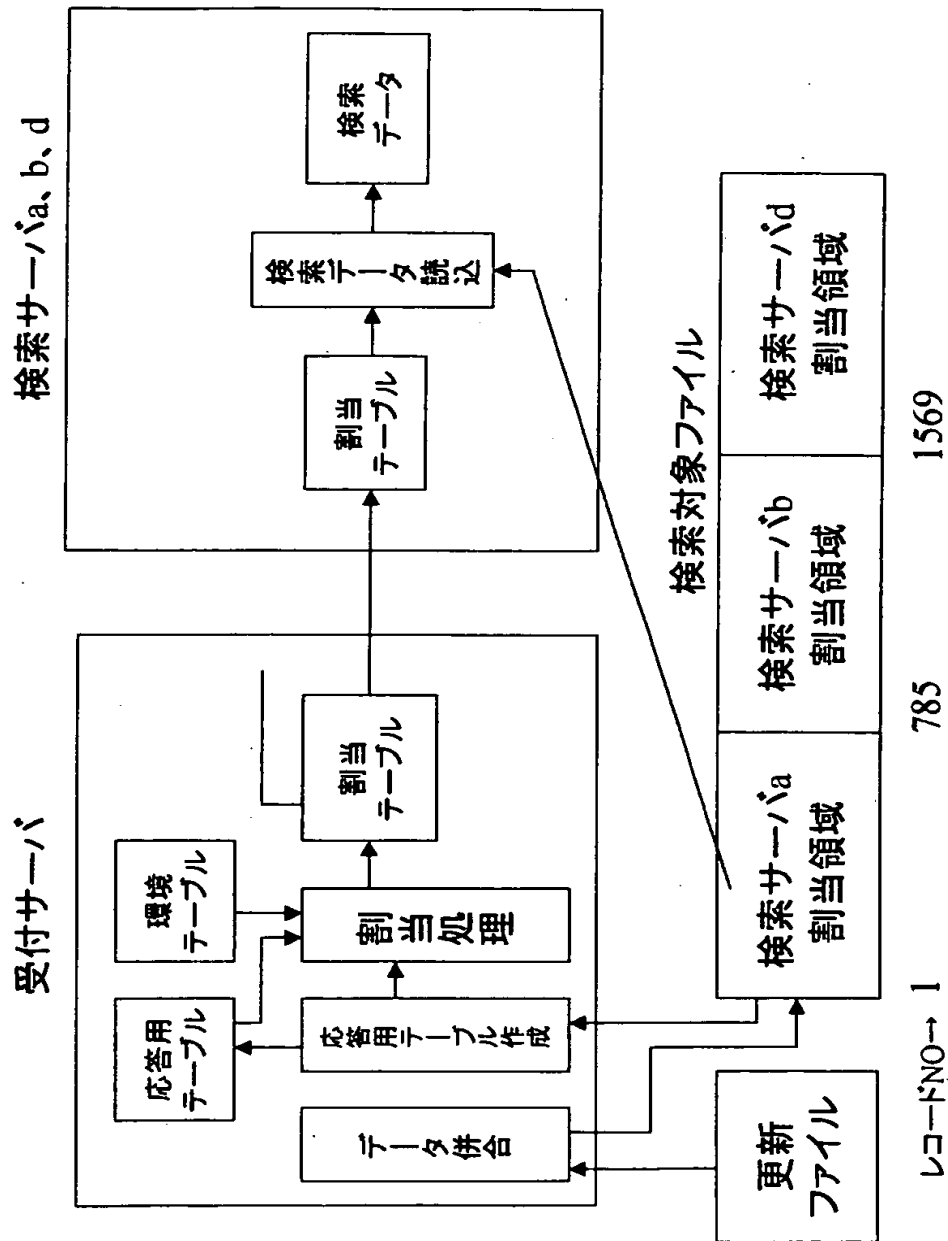
処理の全体の流れ





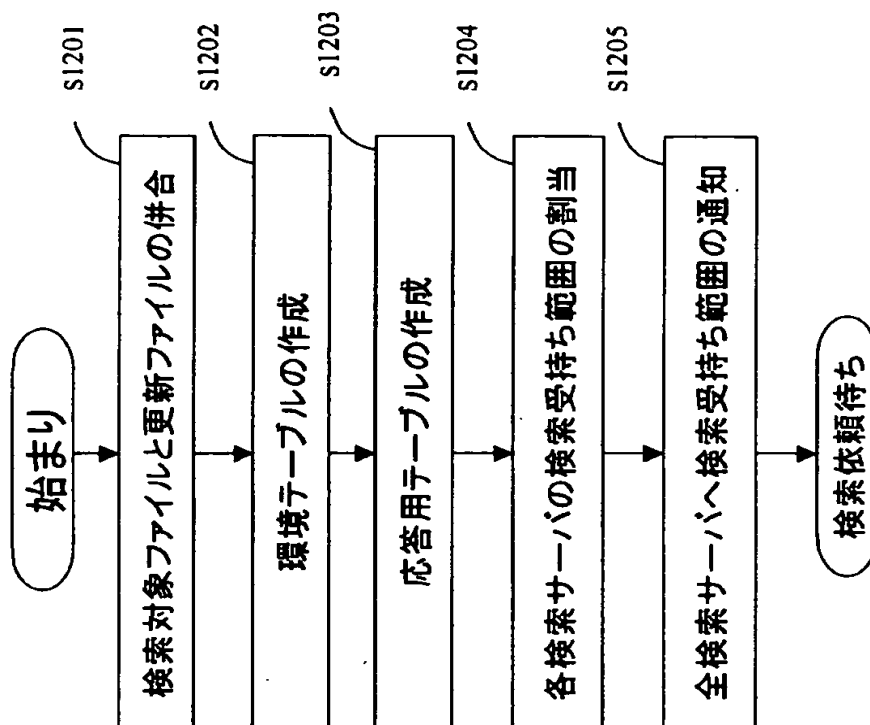
【図 11】

事前処理のブロック図



【図 1 2】

ディレクタプログラムの流れ図(1)



【図 13】

環境テーブルの例(1)

検索サーバ名	アドレス	状態フラグ	更新ファイル 担当フラグ
検索サーバa	15.85.145.01	0	0
検索サーバb	15.85.145.02	0	0
検索サーバc	15.85.145.03	1	0
検索サーバd	15.85.145.04	0	0
検索サーバe	15.85.145.05	0	1

【図 1 4】

応答用テーブルの例

レコードNO	レコードの 先頭アドレス	レコードの 長さ
1	0	250
2	250	580
3	830	322
4	1152	108
.		
.		
.		
2351	452442	242

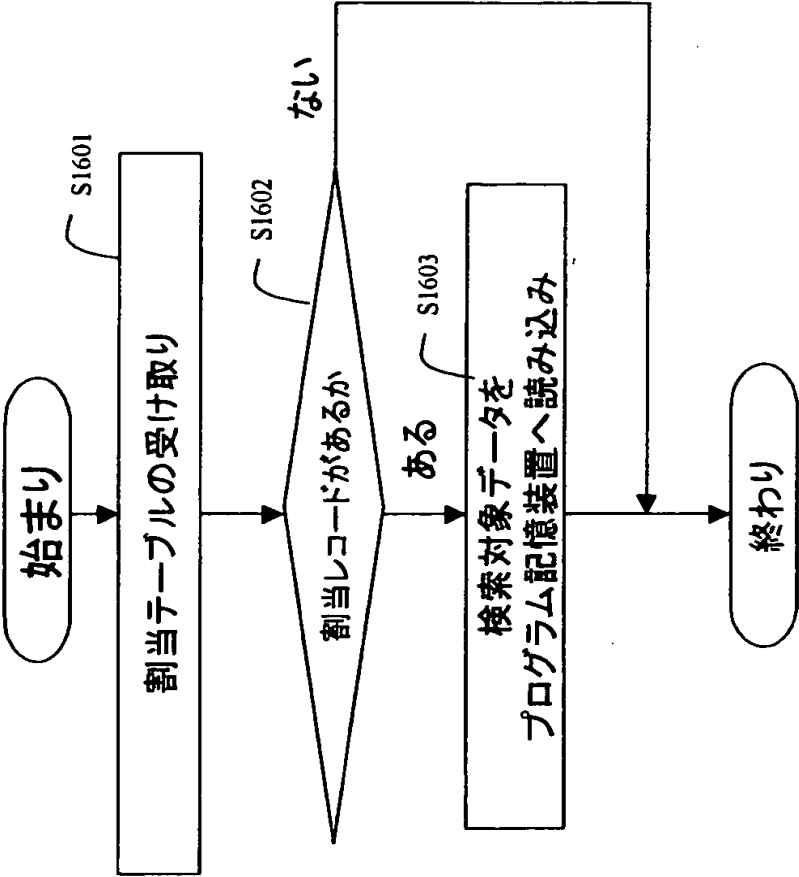
【図 1 5】

検索サーバaの割当テーブルの例(1)

レコードNO	レコードの 先頭アドレス	レコードの 長さ
1	0	250
2	250	580
3	830	322
4	1152	108
.		
.		
.		
784	168997	206

【図 16】

検索サーバの初期設定プログラムの流れ図



【図 17】

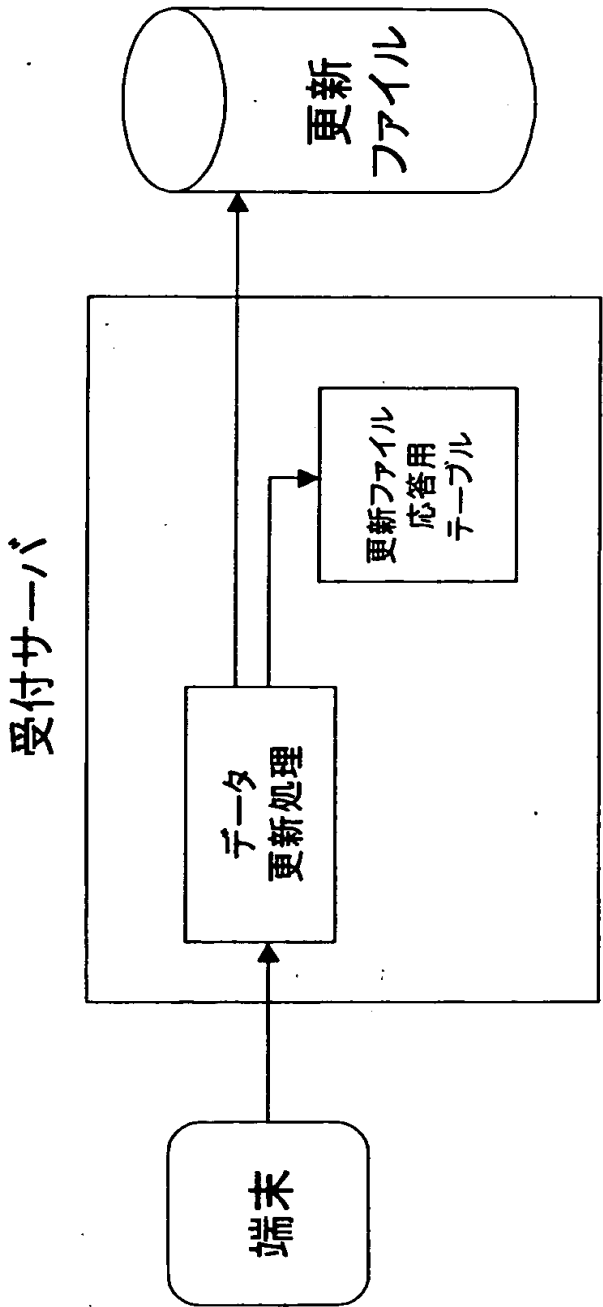
# 検索データの例

[K]<NO>1[K]<名前>松浦正卓[K]<出身校>AA大学[K]<出身地>神奈川県[K]<性別>男[K][R]  
 [K]<NO>2[K]<名前>田端洋子[K]<出身校>BB大学[K]<出身地>石川県[K]<性別>女[K][R]  
 [K]<NO>3[K]<名前>安倍史郎[K]<出身校>CC大学[K]<出身地>大分県[K]<性別>男[K][R]

.  
 .  
 .

【図 1 8】

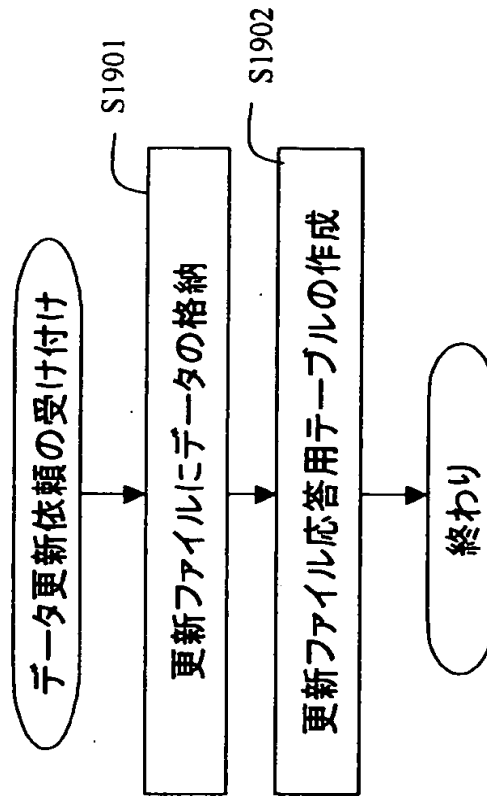
更新処理のブロック図





【図 1 9】

データ更新プログラムの流れ図



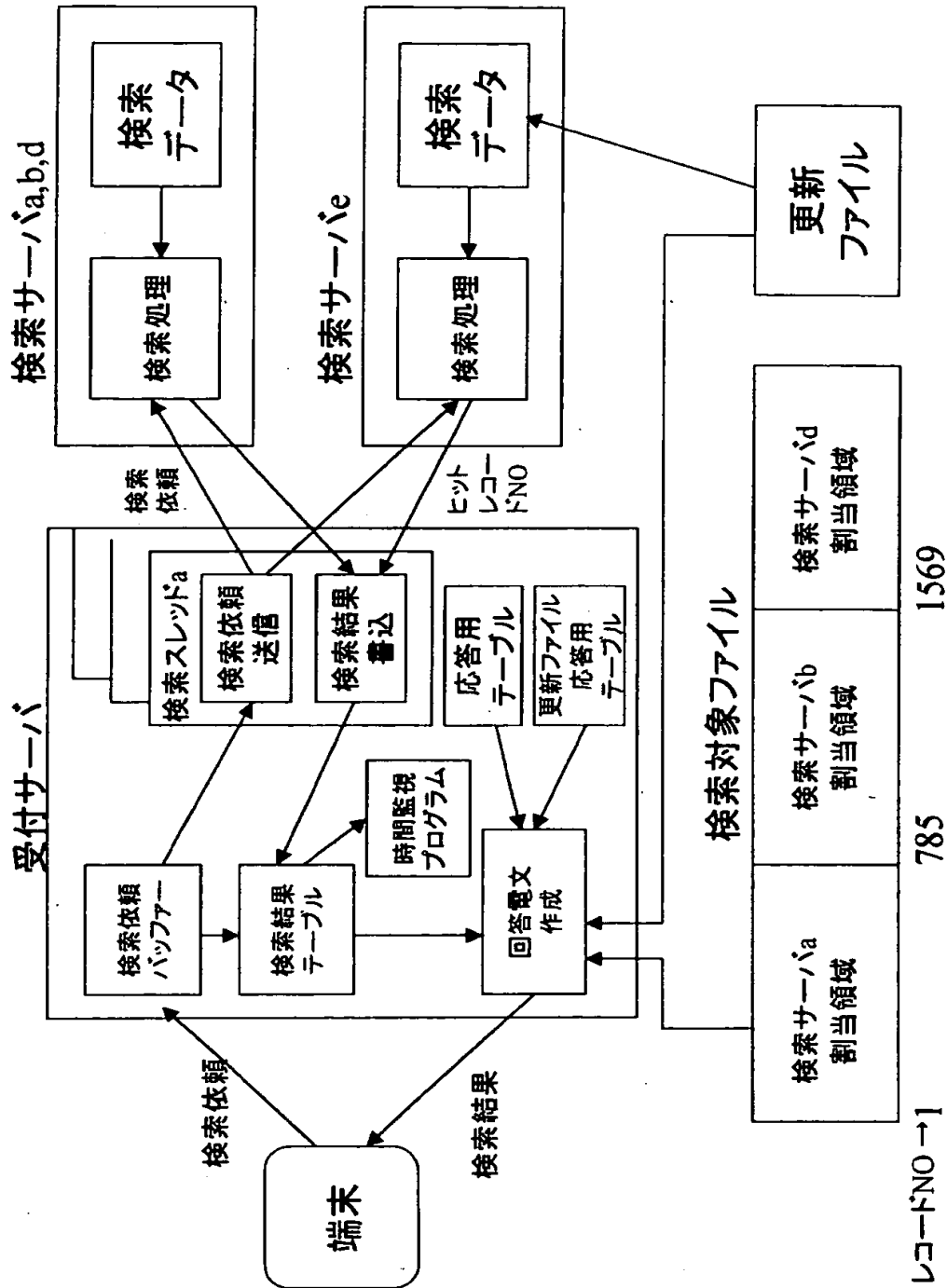
【図 2 0】

更新ファイル応答用テーブルの例

レコードNO	レコードの 先頭アドレス	レコードの 長さ
2352	0	224
2353	224	108
282	332	306
2354	638	278

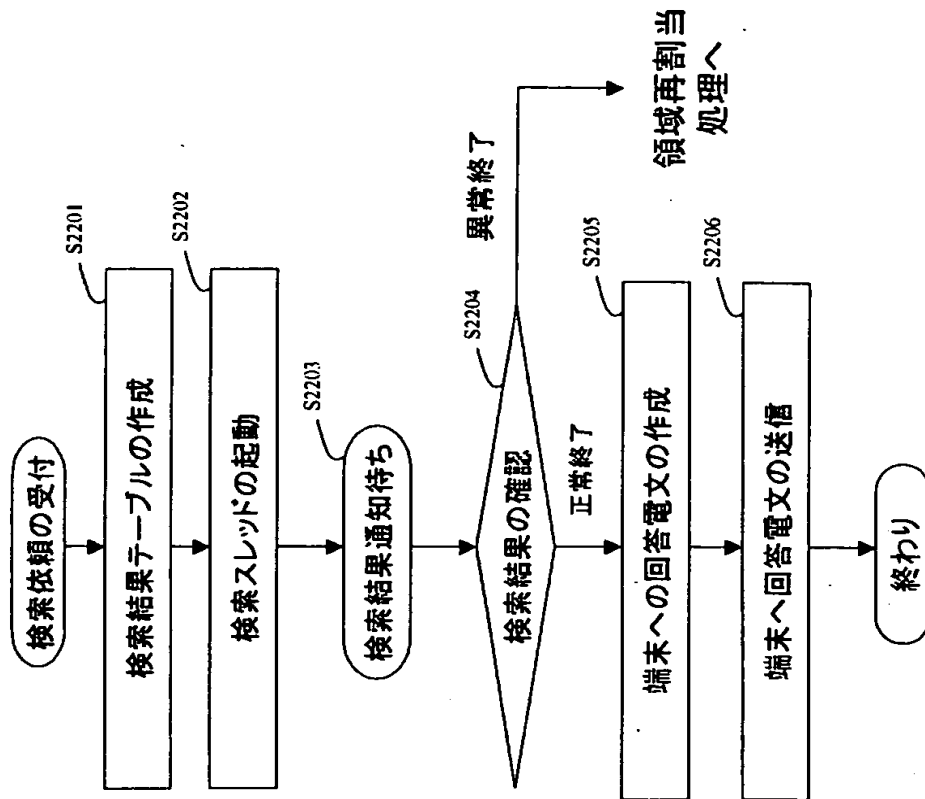
【図 21】

検索処理のブロック図



【図22】

ディレクトリプログラムの流れ図(2)



【図 2 3】

検索依頼バッファー

端末Aの検索依頼
端末Bの検索依頼
端末Cの検索依頼

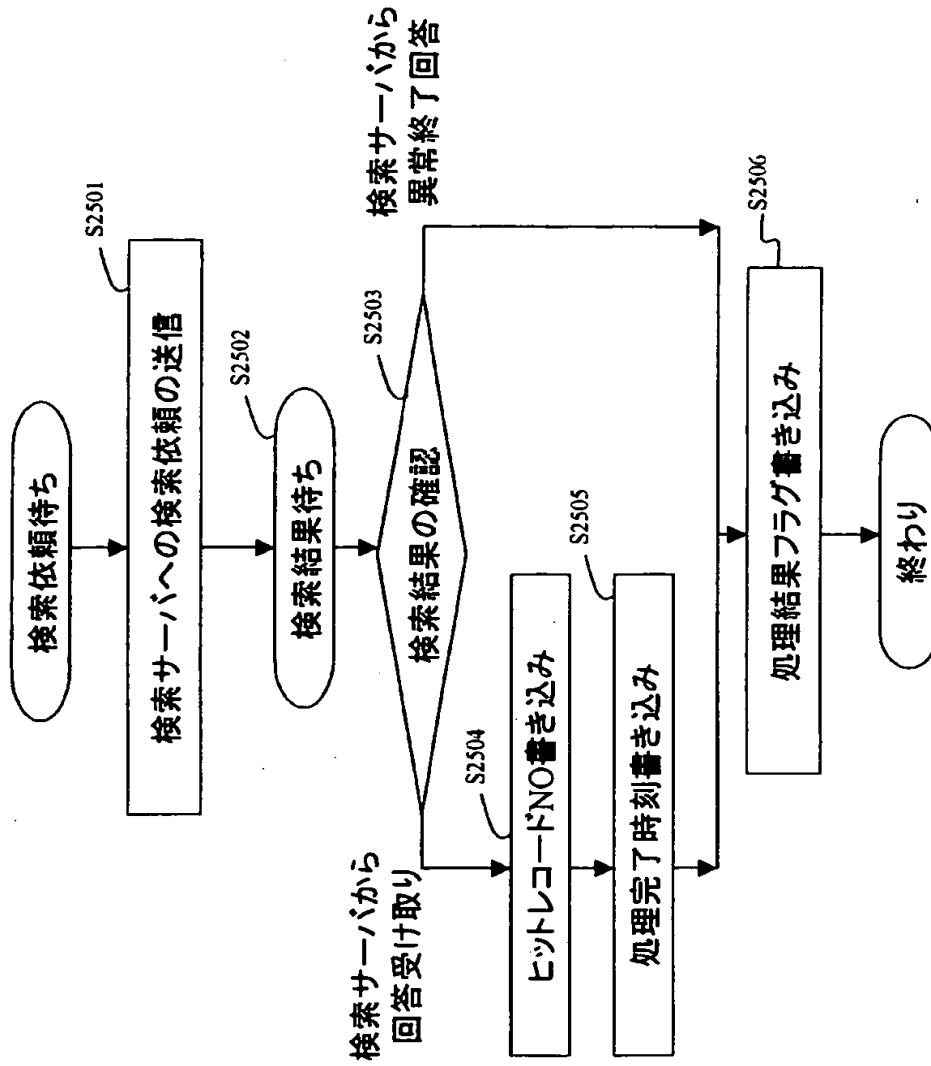
【図 2 4】

検索結果テーブルの例(1)

検索サーバ名	更新ファイル 担当フラグ	処理結果 フラグ	処理完了 時刻	ヒットレコード NO
検索サーバa	0			
検索サーバb	0			
検索サーバd	0			
検索サーバe	1			

【図 25】

検索スレッドの流れ図



【図 2 6】

検索結果の例(1)

ヒットレコードNO	2, 245, 501
-----------	-------------



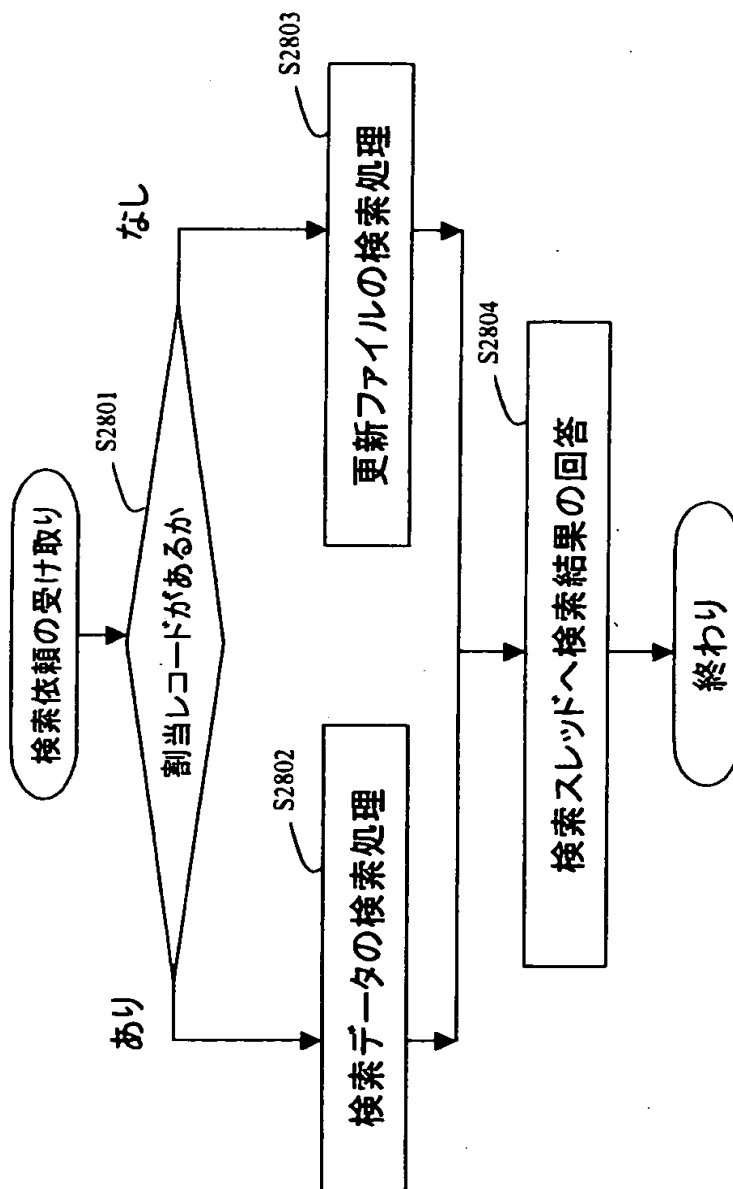
【図 2 7】

検索結果テーブルの例(2)

検索サーバ名	更新ファイル 担当フラグ	処理結果 フラグ	処理完了 時刻	ヒットレコード NO
検索サーバa	0	1	102601541	2、245、501
検索サーバb	0	1	102601545	
検索サーバd	0	2		
検索サーバe	1	1	102601538	2353

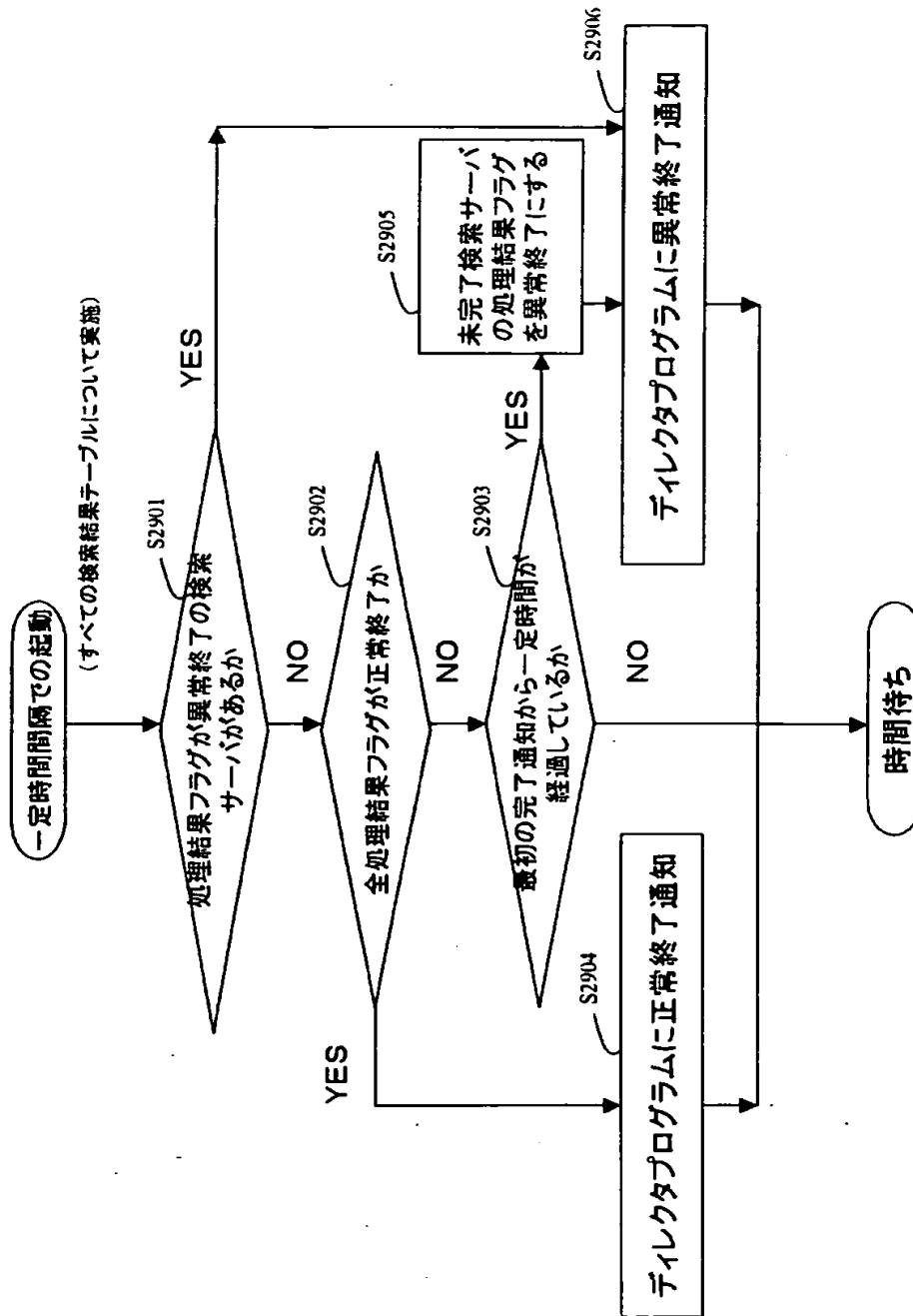
【図 28】

全文検索プログラムの流れ図



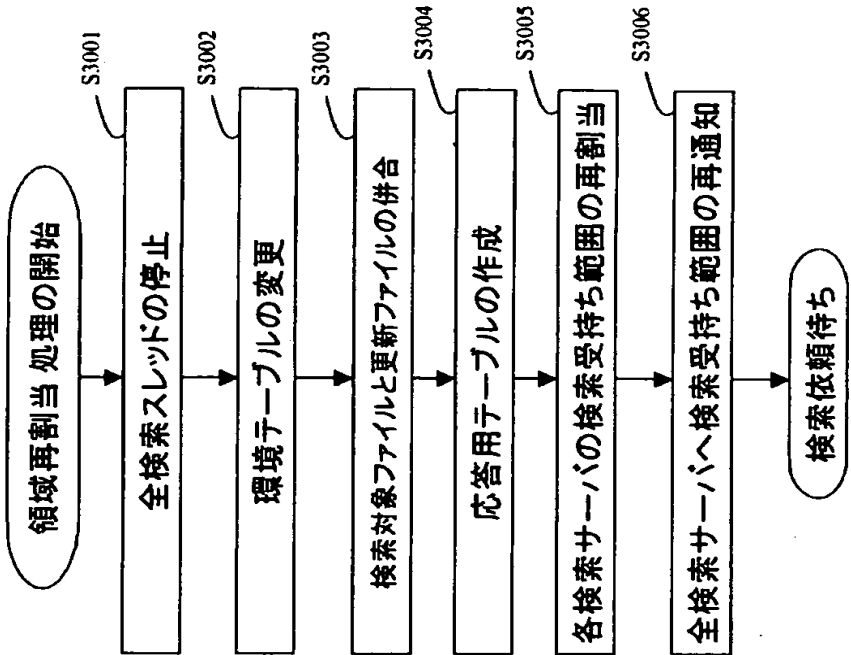
【図29】

時間監視プログラムの処理の流れ



【図 3 0】

ディレクタプログラムの流れ図(3)



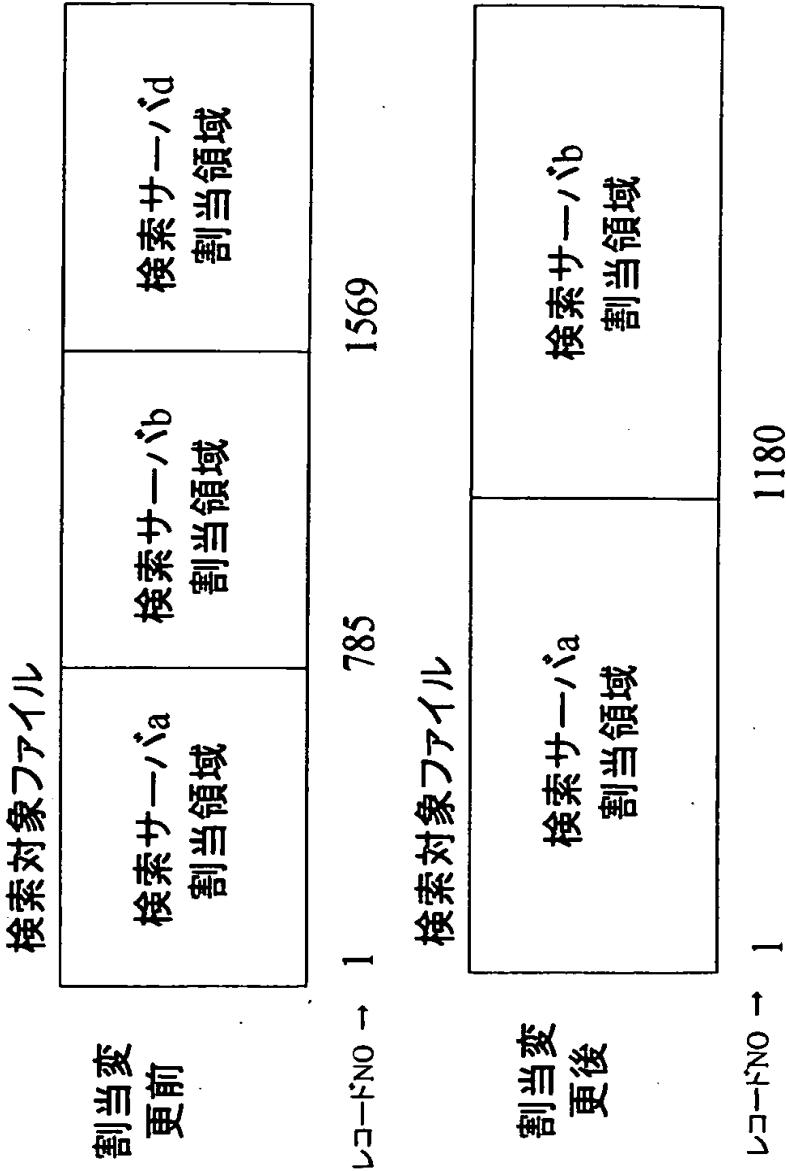
【図31】

環境テーブルの例(2)

検索サーバ名	更新ファイル 担当フラグ	アドレス	状態フラグ
検索サーバ <sup>a</sup>	0	15.85.145.01	0
検索サーバ <sup>b</sup>	0	15.85.145.02	0
検索サーバ <sup>c</sup>	0	15.85.145.03	1
検索サーバ <sup>d</sup>	0	15.85.145.04	1
検索サーバ <sup>e</sup>	1	15.85.145.05	0

【図 3 2】

検索対象領域の再割当



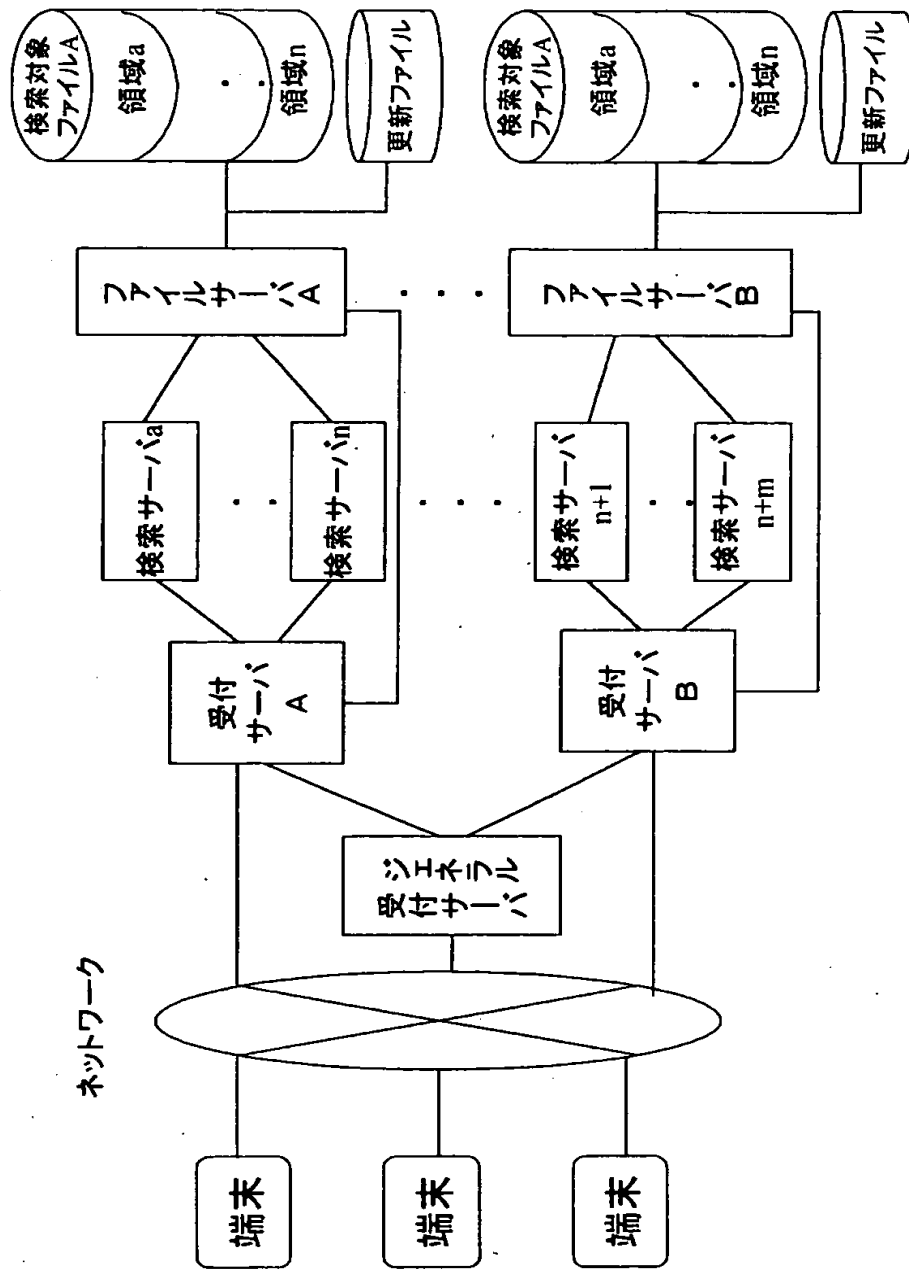
【図 3 3】

検索サーバaの割当テーブルの例(2)

レコードNO	レコードの 先頭アドレス	レコードの 長さ
1	0	250
2	250	580
3	830	322
4	1152	108
.		
.		
.		
1179	246146	188

【図 34】

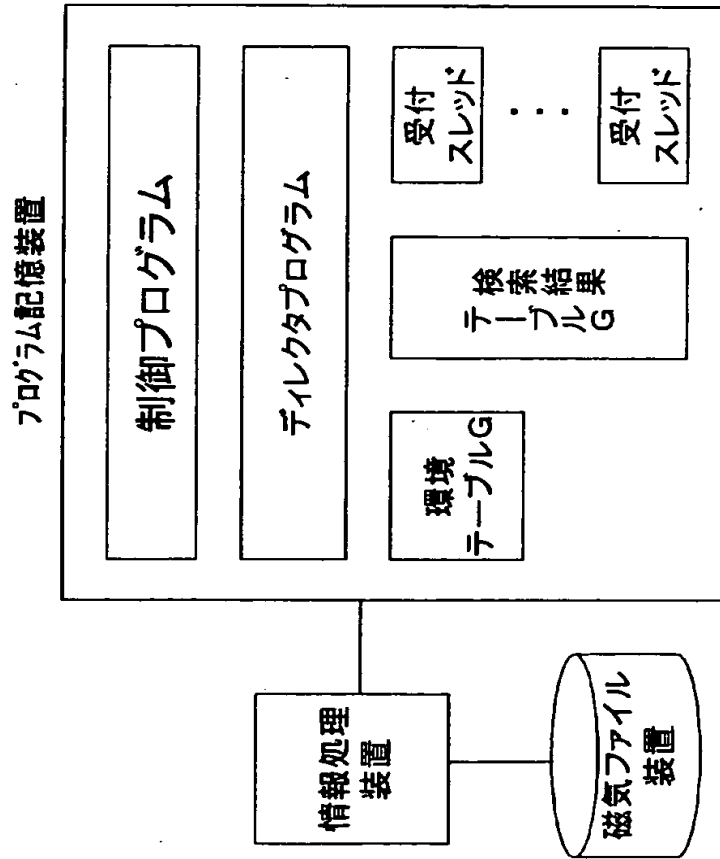
二つ目の実施例の全体構成図





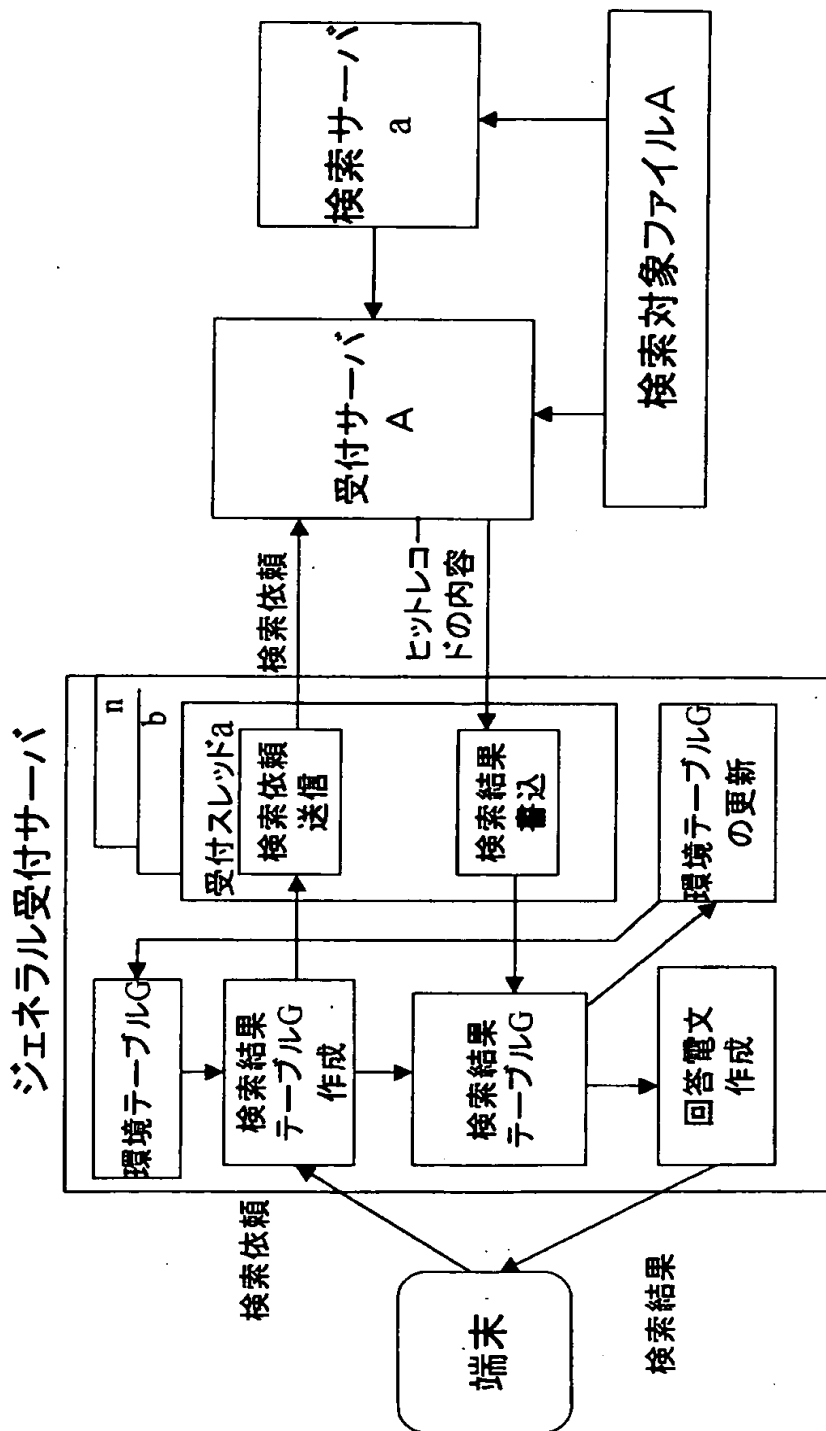
【図 35】

ジェネラル受付サーバの構成図



【図 36】

二つ目の実施例の処理ブロック図



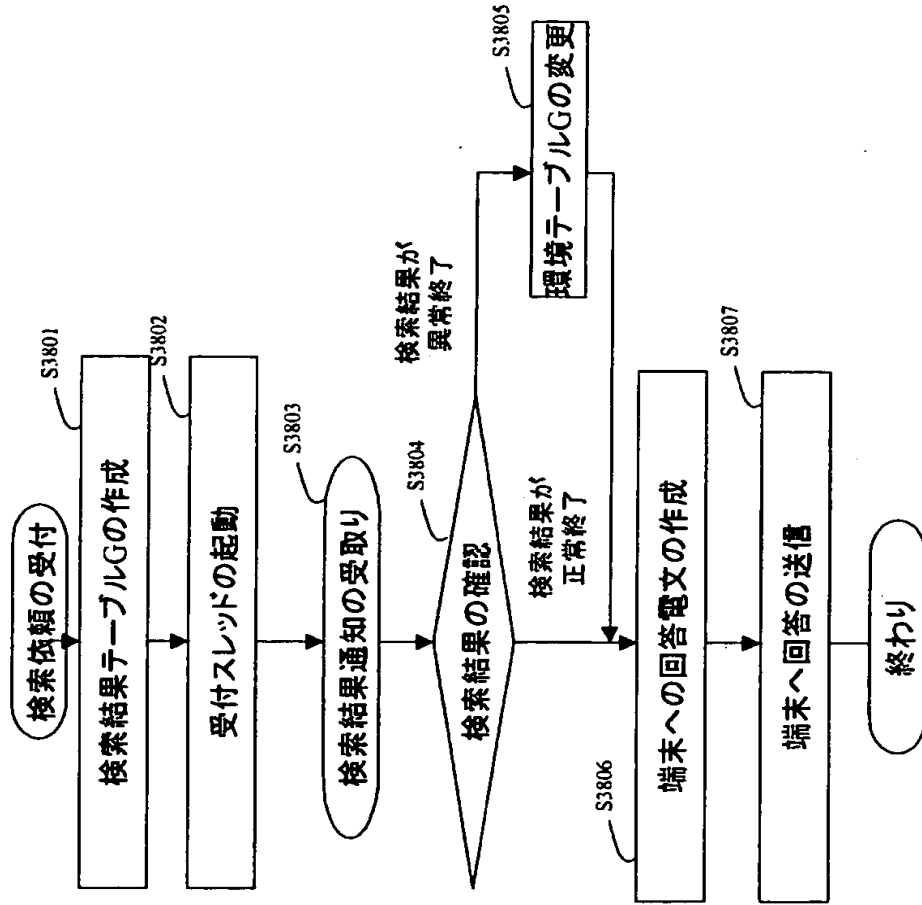
【図 3 7】

## 環境テーブルGの例(1)

受付サーバ名	アドレス	状態フラグ
受付サーバa	16.85.102.01	0
受付サーバb	22.36.155.02	0
受付サーバc	18.24.233.01	0
受付サーバd	34.26.111.01	1
受付サーバe	45.88.188.02	0

【図 38】

ジェネラル受付サーバのディレクタプログラムの流れ図(2)



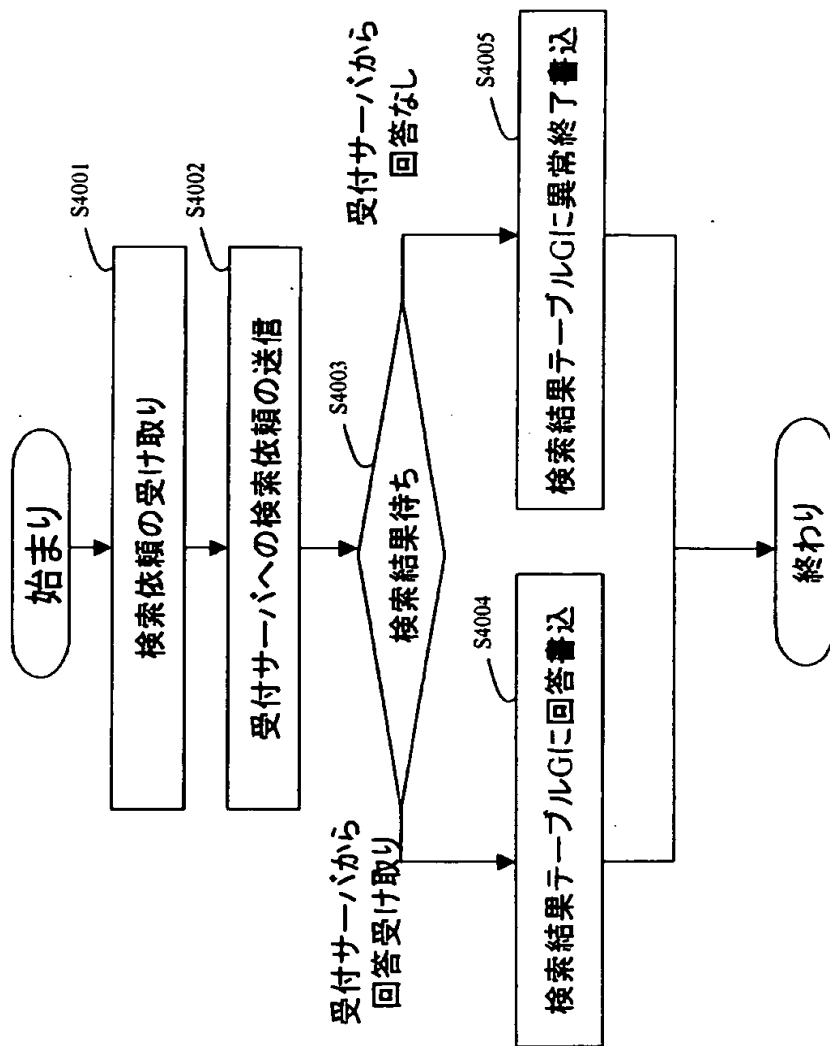
【図 3 9】

検索結果テーブルGの例(1)

受付サーバー名	処理結果 フラグ	検索結果
受付サーバーa		
受付サーバーb		
受付サーバーc		
受付サーバーe		

【図 4 0】

受付スレッドの流れ図



【図 4 1】

検索結果の例(2)

ヒットレコードの内容
[K]<名前>松浦.....,[K]<名前>松浦...

【図 4 2】

検索結果テーブルGの例(2)

受付サーバ名	処理結果 フラグ	検索結果
受付サーバ <sup>a</sup>	1	[ K]<名前>松浦正卓.....
受付サーバ <sup>b</sup>	1	
受付サーバ <sup>c</sup>	2	
受付サーバ <sup>e</sup>	1	[ K]<名前>松浦正卓.....



【図 4 3】

環境テーブルGの例(2)

受付サーバ名	アドレス	状態フラグ
受付サーバa	16.85.102.01	0
受付サーバb	22.36.155.02	0
受付サーバc	18.24.233.01	1
受付サーバd	34.26.111.01	1
受付サーバe	45.88.188.02	0

【書類名】 要約書

【要約】

【課題】 本発明は、複数処理装置を用いて大量データを短時間に検索する文字列照合方式の全文検索システムに関し、検索対象データの更新、追加を検索と並行して行うことが可能で、検索処理装置が故障しても、縮退運転を迅速に行える手段を提供することを課題とする。

【解決手段】 検索対象文字列データ 1 を文字列レコード群に分割して 1 以上の検索処理装置 4 に割り当て、与えられた文字列検索条件を各検索処理装置 4 に送信し、検索結果を受信して統合する検索統合手段 2 を設け、検索対象文字列データ 1 の追加あるいは更新するレコードのみを一時記憶する追加更新一時記憶手段 3 と、あらかじめ定められたいずれか一つの検索処理装置 4 に追加更新一時記憶手段 3 に記憶された文字列レコード群を検索担当データとして指示する追加更新レコード検索指示手段 8 とを備えるように構成する。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000005223]

1. 変更年月日 1996年 3月26日  
[変更理由] 住所変更  
住 所 神奈川県川崎市中原区上小田中4丁目1番1号  
氏 名 富士通株式会社